Dynamic Persuasion and Strategic Search

Yunfei (Jesse) Yao*

(University of California, Berkeley)

July 24, 2023

* Comments welcome. I am indebted to J. Miguel Villas-Boas, Ganesh Iyer, and Yuichiro Kamada for continual guidance and support. I also thank Fred Feinberg, Liang Guo, Sherry He, Kinshuk Jerath, Shachar Kariv, T. Tony Ke, Dmitri Kuksov, Ryota Lijima, Olivia Natan, Ram Rao, Qitian Ren, Chris Shannon, Dong Wei, Pinar Yildirim, Zachary Zhong, Zihao Zhou, Yuting Zhu, and seminar participants at Berkeley Haas, Berkeley Econ theory lunch, Chinese University of Hong Kong, Hong Kong University of Science and Technology, Nova School of Business and Economics, Peking University, University of Illinois Urbana-Champaign, University of Melbourne, University of Rochester, and University of Texas at Dallas for constructive comments. All errors are my own. E-mail address: jesseyao@berkeley.edu.

Abstract

Consumers frequently search for information before making decisions. Since their search and purchase decisions depend on the information environment, firms have a strong incentive to influence it. This paper endogenizes the consumer's information environment from the firm's perspective. We consider a dynamic model where a firm sequentially persuades a consumer to purchase the product. The consumer only wishes to buy the product if it is a good match. The firm designs the information structure. Given the endogenous information environment, the consumer trades off the benefit and cost of information acquisition and decides whether to search for more information. Given the information acquisition strategy of the consumer, the firm trades off the benefit and cost of information provision and determines how much information to provide. This paper characterizes the optimal information structure under a general signal space. We find that the firm only smooths information provision over multiple periods if the consumer is optimistic about the product fit before searching for information. Moreover, if the search cost for the consumer is high, the firm designs the information such that the consumer will be certain that the product is a good match and will purchase it after observing a positive signal. If the search cost is low, the firm provides noisy information such that the consumer will be uncertain about the product fit but will still buy it after observing a positive signal.

1 Introduction

With the rapid proliferation of digital technologies and information channels, it is increasingly common for consumers to seek detailed information before making a decision. More information can lead to less uncertainty and improve decision-making. Since consumers' search and purchase decisions depend on the information environment, firms have a strong incentive to influence it. Advertisers want to choose the advertising content to raise consumers' awareness and interest. Platforms want to design the website to attract traffic. Thus, the consumer faces an endogenous information environment. For example, a consumer considering purchasing a pair of shoes may search on the internet to find out whether or not the item matches his needs.[1] The seller can influence consumers' search and purchase decisions by bidding for search advertising spots to persuade the consumer. Even if the consumer does not purchase after seeing an ad, the seller can keep persuading the consumer through retargeting. The seller decides how much and what kind of information to provide to the consumer. For example, they frequently retarget consumers who have added the product to the shopping cart/list, but only target many other consumers only once. Some sellers show very precise information to consumers, while others add more noise to their ads. At the same time, the consumer spends time and effort searching. He will only search for more information if he anticipates enough gain from it.

The main contribution of this paper is to endogenize the consumer's information environment from the firm's perspective. By considering consumer search and firm information provision simultaneously, we wish to explain why the information environment of consumer search differs across various scenarios. We want to understand when the firm prefers to provide noisy rather than precise information and when it prefers to communicate with consumers for a longer time. We find that the firm provides information incrementally rather than only once if the consumer is optimistic about the product fit before searching for information. If the search cost for the consumer is high, the firm designs the information such that the consumer will be certain that the product is a good match and will purchase it right after observing a positive signal. If the search cost is low, the firm provides noisy information such that the consumer will be uncertain about the product fit but will still buy the product right after observing a positive signal.

Specifically, this paper considers a dynamic model where a receiver (consumer) makes a binary decision between action G (e.g., purchase the product) and B (outside option). There are two states, good (the product is a good match) and bad (the product is a bad match). A sender (firm) always prefers G and sequentially persuades the receiver to take that action. In contrast, the receiver only wishes to take action G if the state is good. Neither the sender nor the receiver knows the state initially but have a common prior belief about it. The receiver can incur costs to search for more information about the state. The updated belief helps him make decisions. If the receiver observes a negative signal, he knows that the state is less likely to be good and will not take action

[1] We refer to the information provider (seller) as "she" and the decision-maker (consumer) as "he" throughout the paper.

G without the arrival of new information. If the receiver observes a positive signal, he knows that the state is more likely to be good, and the expected payoff of taking action G increases.

The sender designs the information structure. Given the endogenous information environment, the receiver trades off the benefit and cost of information acquisition and decides whether to search for more information. The receiver is forward-looking and forms rational expectations of the sender's strategy. The sender can incur higher costs to convince the receiver to take action G with a higher likelihood. Given the information acquisition strategy of the receiver, the sender trades off the benefit and cost of information provision and determines how much information to provide. Therefore, the receiver and the sender simultaneously trade off the benefit and cost of information acquisition/provision.

This paper characterizes the optimal information provision strategy of the sender and the optimal information acquisition strategy of the receiver. There are two periods in the main model. In each period, the sender chooses the information structure, and the receiver chooses whether to search for more information or make a decision. We later extend the two-period model to an infinite-period model and show that the main insights extend to the richer model. Instead of looking at specific parameters of the search environment, we study the design of the information environment under general signal space and characterize the optimal information structure among all feasible information policies. We develop a constrained non-linear programming method to solve the sender's information design problem, because the widely-used concavification method due to Kamenica and Gentzkow (2011) cannot solve it when the receiver's particiaption is strategic.

In equilibrium, the sender induces the receiver to take action G immediately upon observing a positive signal. This way, the sender saves the expected search time and does not need to compensate the receiver for a higher expected search cost. The sender extracts all the surplus from the receiver when she provides information in both periods, while she may leave some surpluses to the receiver when she provides information in only one period. Information smoothing can also save the persuasion cost. So, the sender has an incentive to spread information provision over multiple periods. However, the longer expected search time of the receiver will discourage him from searching if the likelihood of getting a positive ex-post payoff is low. Hence, the sender only smooths information provision if the prior belief is high. If the search cost for the receiver is high, the sender designs the information such that the receiver will be certain that the state is good and will take action G right after observing a positive signal. If the search cost is low, the sender provides noisy information such that the receiver will be uncertain about the state but still takes action G right after observing a positive signal.

We compare the profit-maximizing structure with the efficient (social welfare maximizing) information structure. When the search cost is high, the optimal strategy of the sender also maximizes social welfare because both the sender and the social planner choose the minimal amount of persuasion that can induce the receiver to search. When the search cost is lower, the two information structures are different because the sender does not internalize the receiver's welfare.

In the online appendex, we also consider the implications of discounting and extend the two-period model to an infinite-period model. The main insights extend to those richer models.

1.1 Related Literature

There is a large stream of literature on optimal information acquisition. In particular, consumer search has raised growing interest both theoretically (Stigler 1961, Weitzman 1979, Wolinsky 1986, Moscarini and Smith 2001, Branco et al. 2012, Ke et al. 2016, Ke and Villas-Boas 2019, and Jerath and Ren 2022) and empirically (Hong and Shum 2006, Kim et al. 2010, 2017, Seiler 2013, Honka 2014, Ma 2016, Chen and Yao 2017, Honka and Chintagunta 2017, Seiler and Pinna 2017, Ursu et al. 2020, Moraga-González and Wildenbeest 2021, Morozov 2021, Morozov et al. 2021, and Yavorsky et al. 2021). In the above papers, the information environment is exogenous. Several papers study consumers' endogenous information acquisition. The consumer chooses both the search rule and the information environment. In Zhong (2022a), the decision-maker gradually gathers information about one product. Poisson learning is optimal for him. In Guo (2021), the consumer sequentially search for information about multiple products and determines how much to evaluate each product. Because the strategy and the outcome depend on the information structure, other payoff-relevant parties (e.g., firms) have a strong incentive to influence it. We take this into account by having the firm endogenously determine the information environment of the consumer.

Some papers investigate the design of the search environment from the firm's perspective. In Dukes and Liu (2016) and Kuksov and Zia (2021), the platform or the seller select the search cost to influence the consumer's search strategy. In Villas-Boas (2009), Liu and Dukes (2013), and Kuksov and Lin (2017), the product line design of the seller impacts consumer's search decision. Villas-Boas and Yao (2021) consider the optimal retargeting strategy of the firm which advertises to consumers who have a high likelihood of considering the firm's product. By advertising, the firm increases the frequency of consumers' learning information and the ability to track consumers. In Zhong (2022b), the platform recommends relevant sellers based on match values and prices to consumers. The platform designs the search algorithm by picking the match precision and the relative importance between prices and match values. Comparing the welfare outcomes among information structures emphasizing different vehicle characteristics under the counterfactuals of their structural model, Gardete and Hunter (2020) find that emphasizing the vehicle's history and obfuscating price information improves both consumer and firm welfare. Mayzlin and Shin (2011) consider a setting where the consumer can obtain an exogenously given signal by searching for information about the product quality. They find that uninformative advertising may serve as an invitation for the consumer to search. In Yao (2023), the firm can affect consumers' belief through informative advertising before they engage in costly search. In related literature on choice overload, the seller determines the amount of information provided to the consumer (Kuksov and Villas-Boas 2010, Branco et al. 2016). The decision of the seller affects the consumer's search cost.

While the above papers are related to our paper, there are substantial differences. Instead of looking at specific parameters of the search environment, we study the design of the information environment under a general signal space. We characterize the optimal information structure among all feasible information policies. In the existing literature, the consumer either fully observes the value of a product/attribute or gets a noisy signal from an exogenously specified distribution. In contrast, in our paper, the firm determines the distribution of the signal. The optimal information structure can be asymmetric and may not correspond to standard distributions.

We use a belief-based method of modeling information provision, first introduced by Aumann and Maschler (1995) and Kamenica and Gentzkow (2011) in the Bayesian persuasion and information design literature.[2] The sender picks a mean-preserving spread of the prior belief as the posterior belief, which simplifies the analysis. Some papers have studied the persuasion problem where either the receiver or the sender incurs costs. For example, Ball and Espín-Sánchez (2022) study a persuasion problem in which the sender chooses from a restricted set of feasible experiments, and the experiment can be costly. In Degan and Li (2021), the sender's persuasion cost depends on the precision of the signal. Wei (2021) considers a static persuasion problem in which a rationally inattentive receiver incurs information processing costs. Jerath and Ren (2021) consider a static model in which the consumer chooses the optimal information structures, taking into account that he needs to incur a cost to search for and process the signals. Instead of directly providing information, the firm influences the consumer's information environment by imposing constraints on the precision of the signals. Berman et al. (2022) study the information design of the recommendation algorithms under endogenous pricing and competition. Gentzkow and Kamenica (2014) extend the widely-used concavification approach of Kamenica and Gentzkow (2011) to the setting where the sender's cost is posterior-separable.

We contribute to this literature by allowing the receiver to search for information voluntarily. The concavification approach cannot be used if the receiver's participation is strategic. So, we instead develop a constrained non-linear programming method to solve the sender's information design problem in the presence of consumer strategic search. Because we consider a dynamic problem, different information structures may correspond to different forms of sender's objectives and receiver's participation constraints. To reduce the dimensionality of the problem, we first show that the optimal information structure must induce the receiver to take action G immediately after receiving a positive signal. This qualitative property greatly simplifies the problem as we can limit our attention to such information structures. Nevertheless, the constraints are non-linear and consist of multiple variables, making the optimization problem challenging. To make the problem tractable, we transform the program into a set of constrained programs, each with one constrained variable. We then select the global solution by comparing the local solutions to each program.

Ke et al. (2022) study how online platforms should design the information in the presence of consumer search. In their paper, the information impact both consumer search and targeted advertising and allow them to study the trade-off between sales commission and advertising revenue.

[2] See Bergemann and Morris (2019) for a survey of this literature.

The firm designs the information to manipulate consumer's belief prior to search. The consumer always process this information but can search for more information strategically given an exogenous information structure (full revelation upon search). Our contribution is to integrate the information provision with consumer search and fully endogenize the search environment.

While some recent papers extend the static setting of Bayesian persuasion to the dynamic one, the persuasion cost is usually zero or a constant (Ely 2017, Renault et al. 2017, Ball 2019, Che et al. 2020, Ely and Szydlowski 2020, Orlov et al. 2020, Iyer and Zhong 2021, Bizzotto et al. 2021). In our model, the sender incurs a persuasion cost, and the receiver incurs a search cost. Unlike most Bayesian persuasion literature, the receiver's participation is strategic. The sender may need to convince the receiver to search or speed up the receiver's learning by incurring a higher cost. Therefore, we can investigate the sender's optimal trade-off between the benefit and cost of information provision.

The remainder of the paper is organized as follows. Section 2 presents the main model. Section 3 introduces and analyzes some benchmarks. Section 4 characterizes the optimal information provision strategy and the equilibrium outcomes of our problem. Section 5 characterizes the efficient information provision strategy and summarizes the information distortion when the sender rather than the social planner designs the information structure. Section 6 concludes.

2 The Model

2.1 States, Actions, and Payoffs

There are two players, a sender and a receiver, and two states, good (g) and bad (b). The receiver ultimately makes a binary decision between G and B. The sender wishes to persuade the receiver to take action G regardless of the state, while the receiver wishes to match the decision with the state (taking action $G(B)$ when the state is $g(b)$). There is no discounting. The payoffs of the decision for the players are the following:

(sender payoff, receiver payoff)	action G	action B
state g	(p, v_g)	$(0, 0)$
state b	(p, v_b)	$(0, 0)$

The sender earns a positive payoff, $p > 0$, if the receiver takes action G. The receiver's payoff is positive if he takes action G when the state is g, $v_g > 0$, and negative if he takes action G when the state is b, $v_b < 0$. Both players get zero payoff if the receiver takes the action B (which can be thought of as an outside option). We assume without loss of generality that $v_g = 1 + v_b$.[3] Neither the sender nor the receiver knows the state initially but have a common prior belief about it, $\mu_0 := \mathbb{P}(\text{the state is } g) \in (0, 1)$. It summarizes all information the receiver has before searching

[3] This assumption is without loss of generality because we can normalize the payoffs by $v_g - v_b$.

for the sender's information. In each period $t \in \{0, 1\}$, the sender determines and commits to the information structure of the current period but cannot commit to the information structure in the future. The receiver can search for information (action S) before deciding. The information acquisition is costly but helps the receiver make better decisions. If the receiver chooses to search, he incurs a search cost c and observes the realization of a binary signal $s \in \{0, 1\}$ that reveals some information about the state. By choosing $\mathbb{P}[s = 1|g]$ and $\mathbb{P}[s = 1|b]$, the sender uniquely determines the signal. We order the value of the signal such that $\mathbb{P}[s = 1|g] > \mathbb{P}[s = 1|b]$. Hence, $s = 1$ corresponds to a positive signal and $s = 0$ corresponds to a negative signal. Players update the belief about the state according to Bayes' rule after the realization of the signal.[4] The game ends whenever the receiver makes a decision (G or B). Figure 1 illustrates the timing of the game.

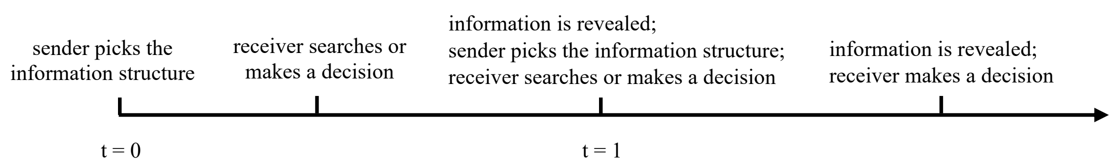

Figure 1: Timing of the Game

Analogous to Proposition 1 of Kamenica and Gentzkow (2011), we can work with mean-preserving posterior beliefs rather than the specific signal structure to simplify the analysis. Specifically, the existence of a binary signal is equivalent to the existence of a binary-valued posterior belief whose expectation is equal to the prior belief. We state this result and its proof formally in the online appendix. Denote the belief at the beginning of each period by μ_t. In each period, with probability λ_t, the receiver observes a positive signal and the belief increases to $\bar{\mu}_t$. We refer to λ_t as the probability of a positive signal and $\bar{\mu}_t$ as the belief after observing a positive signal. With probability $1 - \lambda_t$, the receiver observes a negative signal, and the belief decreases to $\underline{\mu_t}$. We refer to $\underline{\mu_t}$ as the belief after observing a negative signal.

Since the sender designs and provides information to the receiver while the receiver does not need to collect the information, different information should cost differently for the sender but not for the receiver. Therefore, we assume that the receiver incurs a flow cost of c per period of search, while the sender's cost of information provision is increasing and convex in the probability of a positive signal, $K = K(\lambda)$. It is relatively cheap for her to provide information with a low λ. The marginal cost increases at an increasing rate as λ increases.

Assumption $K(\cdot) \in \mathcal{C}^2(\mathbb{R}_+), K'(\lambda) > 0, K''(\lambda) > 0, K(0) = 0, \lim_{\lambda \to 1^-} K'(\lambda) = +\infty, \lim_{\lambda \to 0^+} K'(\lambda) = 0.$

Throughout this paper, we refer to the likelihood of a positive signal as the *amount of persuasion*. The total payoff for each player is the payoff of the decision net of the information

[4] We can assume without loss of generality that the sender also observes the signal realization. This is because the sender can perfectly infer the signal realization from the receiver's action under the optimal signal structure, according to Proposition 1.

provision/acquisition costs. The receiver is forward-looking and forms rational expectations about the sender's strategy in the future. To avoid the trivial case in which the sender provides no information and the receiver always takes the sender's desired action, we assume that $\mu_0 v_g + (1-\mu_0)v_b < 0 \Leftrightarrow \mu_0 < -v_b$. So, the receiver will never take action G without searching. We also assume that the search cost is not too high, $c < v_g$. Otherwise, the receiver will never search.

Sender's Convex Cost

To persuade the receiver to take action G, the sender wants to increase the receiver's belief about the good state. It is easy to provide information that increases the receiver's belief with a low probability but hard to do so with a high probability, and impossible to always increase the receiver's belief. So, we assume that the sender's cost of information provision is increasing and convex in the amount of persuasion. The convex information provision cost is common in the literature (e.g., Robert and Stahl 1993). In section 2.2, we connect our model to three real-world examples and discuss why the sender's cost may be convex. In section 3.2, we work out an alternative model with a constant information provision cost and show that the convexity of the sender's cost plays a critical role when the sender persuades the receiver gradually. However, it neither means that the sender's cost always has this particular form nor that we need to take this assumption literally. Other cost functions, such as constant cost, may be more appropriate in other settings. Especially in settings where the sender persuades the receiver only once, a constant cost may be the driven force. But, some level of convexity in the amount of persuasion is essential when the persuasion occurs in more than one period. In dynamic persuasion problems, our assumption captures the sender's incentive to smooth information in a simple way and provides new insights.

It is also reasonable to expect that the sender's cost depends on the true state. For example, it may be more costly for her to persuade the receiver to take action G when the state is b. Since the sender does not know the state and thus her true cost, she can only use her expected cost in decision-making. The problem becomes intractble. However, as we can see in the next section, the underlying mechanisms will not be qualitatively affected.[5]

Receiver's Information Acquisition

We consider the receiver's search cost but not his attention cost, commonly modeled in the rational inattention literature. Let's first distinguish the search cost and the attention cost. The receiver incurs search costs to observe the realization of the signal while incurs attention costs to process/interpret the signal. The receiver needs first to observe the signal to process it. So, search cost is the prerequisite of attention cost. In the standard consumer search literature, people only

[5] An alternative tractable way of incorporating the state-dependent cost into our model is to allow for product return if we interpret our model as consumer search before product sales. The buyer can return the product after purchasing it and realize it is a bad match. The seller will incur a re-stocking cost upon return. So, her cost of persuading the buyer to buy a bad product is higher than her cost of persuading the buyer to buy a good product. When it is costless to return the product, the buyer will always buy without searching and return the product if it turns out to be a bad match. However, even though the buyer can get a full refund from many marketplaces such as Amazon, he needs to incur time and effort to bring the package to the store or a shipping carrier. The qualitative properties of the main model still hold as long as a proportion of the buyers have enough return costs.

model search costs. Adding another layer of receiver attention costs will make the model intractable and distort us from the sender's information design problem. However, it can be interesting in settings where the receiver needs to incur a lot of effort to process the signal.

Suppose we add attention cost to our model. Since the receiver has to incur an additional cost to process the signal, the sender must give him more benefit from searching. Otherwise, the receiver will not search. So, the sender's persuasion cost will increase, and the information distortion will decrease. The sender will also have less incentive to smooth information over two periods.

Symmetric Information

The sender and receiver have symmetric information in Bayesian persuasion problems. It is important to note that the receiver can gather information from a different channel, which will be incorporated into the prior belief, and the sender does not need to report everything she knows. Data analytics helps the sender predict the receiver's prior using detailed demographic and behavioral data. Symmetric information means the sender can commit to any information structure she designs. She cannot change the signal realization after observing it. So, the sender and the receiver update their belief in the same way after the sender commits to an information structure.

The Commitment Assumption

This standard assumption in the Bayesian Persuasion literature is less restrictive than it may seem, as discussed in detail in the Section I. C of Kamenica and Gentzkow (2011). In particular, it can be mandated by legal reasons (e.g., many laws about the advertising content, clinical trials need to be pre-registered) or through reputation (e.g., consumers can observe the frequency of positive signals through repeated interaction with the firm).

2.2 Applications

We can apply our model to many settings where a sender wishes to persuade a receiver to take a particular action while the receiver wishes to match the action with the state and can strategically gather more information before making the decision. We discuss three applications of the model.

Value-added Branding in Luxury Goods

A luxury goods company wants to add value to its branding. In each period, the seller (sender) can spend money on building the luxury image of her brand. An example of the spending is advertising, which has increasing marginal costs. The prior comes from country-of-origin effects: a new brand of Swiss watches has a high prior compared to a new brand of Polish watches, and vice versa for vodka. So Polish watchmakers and Swiss vodka distillers only have one shot at persuading the consumer (receiver) to buy, while Polish vodka distillers and Swiss watchmakers get a second chance. The consumer's cost of receiving signals is c. It may be more costly for the seller to reach richer and busier customers. Our model allows for any correlation between the sender's and receiver's costs.[6]

[6] We thank the AE for suggesting this example.

Online Advertising

Consumers (the receiver) frequently search online to learn more about the product before making a purchasing decision. The advertiser (sender) can target a consumer (showing one ad) or retarget him (showing multiple ads to the same consumer). In each period, the advertiser can bid for the advertising spot and persuade the consumer to purchase the product. She directly controls how precise the advertising content is but can only affect the winning probability indirectly through the bidding amount. The likelihood that the consumer receives a positive signal (λ) is proportional to the winning probability, given the advertising content. In a symmetric N-bidder first-price auction with i.i.d. uniform distribution of the valuation, the bidding amount is convex in the winning probability. The consumer's cost is their time. The prior belief can come from word of mouth or past experience. Advertisers most frequently retarget consumers who have added the product to the shopping cart/list. Our model can explain this because those consumers, on average, have a high prior belief about the product. We will show that the sender wants to provide information in both periods if the prior is high while only wants to provide information in one period if the prior is low.

Lobbying

Companies (sender) have a strong incentive to influence a regulator's (receiver) policymaking. The regulator is deciding between policies A and B, which can be a high or low standard on car/food safety, a strict or relaxed policy on privacy, etc. The company wants to persuade the regulator to choose policy A. It can show some information to the regulator. The information can be anything that may affect the regulator's decision-making and does not need to be about the company's own product. The regulator's prior belief comes from any information from a different source. λ is the likelihood that the regulator likes more about policy A after the company lobbies. The company's cost is the effort it incurs to persuade the regulator. Raising the effort level has an increasing marginal cost. The regulator's cost is the opportunity cost. If many other companies have valuable information to present, the regulator will have a higher search cost. If the regulator is leaning toward policy A before the company lobbies, the company has two chances to convince the regulator to "follow his heart". If the regulator is leaning toward the opposite policy, policy B, the company only has one shot. The regulator will not listen to it again if the information in the first period makes him like policy A even less.

2.3 Strategies and Equilibrium Concepts

Since the belief is common knowledge and there is no private information, we consider the sender-preferred subgame perfect equilibrium, as in Kamenica and Gentzkow (2011). If multiple actions (B, G, and S) give the receiver the same expected payoff, we assume that the receiver chooses an action that maximizes the sender's expected payoff. We also assume that the sender prefers to give the receiver more surplus in the first period if more than one equilibrium lead to the same expected sender's payoff. A perturbation of very little discounting justifies this assumption.

2.3.1 Equilibrium in the Second Period

The receiver has to make a decision between G and B at the end of the second period. Since it is costly for the sender to provide information, the sender will either give no information or provide information such that the receiver searches and takes action G if a positive signal arrives. We illustrate the belief evolution in the second period in Figure 2. Also, the distribution of the belief induced by the signal should be a mean-preserving spread of the initial belief: $\mathbb{E}[\Delta\mu] = 0$. In sum, the sender either does not provide information and obtains zero payoffs or takes into account the following constraints when designing the information structure:

(1) participation constraint:

$$\lambda_1[\bar{\mu_1} v_g + (1-\bar{\mu_1})v_b] = \lambda_1(\bar{\mu_1} + v_b) \geq c \tag{IR_1}$$

Notice that the receiver will take action B if the signal realization is negative because his belief decreases. If he takes action G after either a positive or a negative signal, there is no benefit from searching in the second period and he will not search.

(2) feasibility constraint:

$$\lambda_1\bar{\mu_1} + (1-\lambda_1)\underline{\mu}_1 = \mu_1 \tag{F_1}$$

If the sender provides information, the constrained program of the sender is:

$$\max_{\lambda_1,\bar{\mu_1}} -K(\lambda_1) + p\lambda_1 \tag{P_1}$$
$$\text{s.t. } (IR_1), (F_1), \lambda_1 \in [0,1], \underline{\mu}_1 \in [0, \mu_1)$$

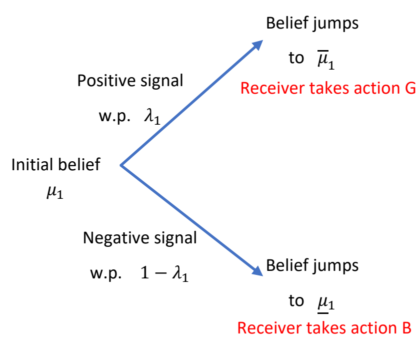

Figure 2: Belief Evolution in the Second Period

We analyze the solution to this problem in section 4. Though the information structure consists of $(\lambda_1, \bar{\mu_1}, \underline{\mu}_1)$, any two of them fully characterize the strategy because the third variable is then uniquely determined by (F_1). Therefore, we use $(\lambda_1, \bar{\mu_1})$, the probability of a positive signal and the belief after observing a positive signal, to represent the sender's strategy.

2.3.2 Equilibrium for the Entire Game

The sender cannot control the receiver's action directly, as the receiver searches for information and makes the decision voluntarily. However, by designing different information, the sender can rationally anticipate the receiver's action. In other words, the information structure designed by the sender induces different receiver behaviors.

The sender has three options. Firstly, she can provide no information and obtain zero payoffs.

Secondly, the sender can provide information in only one period. If the receiver decides to search, he will take action G if a positive signal arrives and taks action B if a negative signal comes. We call this type of signal a *one-shot* signal because a single positive signal raises the receiver's belief high enough and suffices to convince the receiver. The sender will not provide extra information regardless of the signal realization. Her problem is exactly (P_1) if we ignore the time subscript. In the base model without discounting, it does not matter whether the sender provides information in the first or the second period. We assume that the sender provides information in the first period in this case for notational simplicity.

Lastly, the sender can provide information in both periods.[7] There are two classes of signals for the sender to choose from. By providing relatively precise information in both periods, the sender induces the receiver to take action G immediately after observing a positive signal. We call this type of signal a pair of *one-shot* signals because a single positive signal in either period raises the receiver's belief high enough and suffices to convince the receiver to take action G. If a negative signal arrives in the first period, the sender will provide another signal, hoping that a positive signal will arrive in the second period. Figure 3 illustrates the belief evolution processes if the sender provides a one-shot signal or a pair of one-shot signals. The receiver's belief increases after observing a positive signal and decreases after observing a negative signal.

By providing less precise information in both periods, the sender induces the receiver to take action G only after searching for information in both periods. Even if the receiver receives a positive signal in the first period, his will still have enough uncertainty and voluntarily search again to learn more about the state. We call this type of signal a pair of *iterative* signals because the sender can only convince the receiver to take action G after the receiver searches for information and updates his belief iteratively in both periods. By providing a pair of iterative signals, the sender encourages the receiver to search more. Figure 4 illustrates the belief evolution processes if the sender provides a pair of iterative signals. If the sender designs information such that the negative signal is really bad and reduces the receiver's belief by a lot, the receiver will keep searching only after observing a positive signal in the first period. If instead, the sender designs information such that the negative signal reduces the receiver's belief by a little, the receiver will keep searching regardless of the signal realization in the first period.

[7] By providing information in both periods, we mean that the sender will provide information in the second period under some circumstances. If the receiver makes a decision in the first period, the sender clearly will not provide information in the second period.

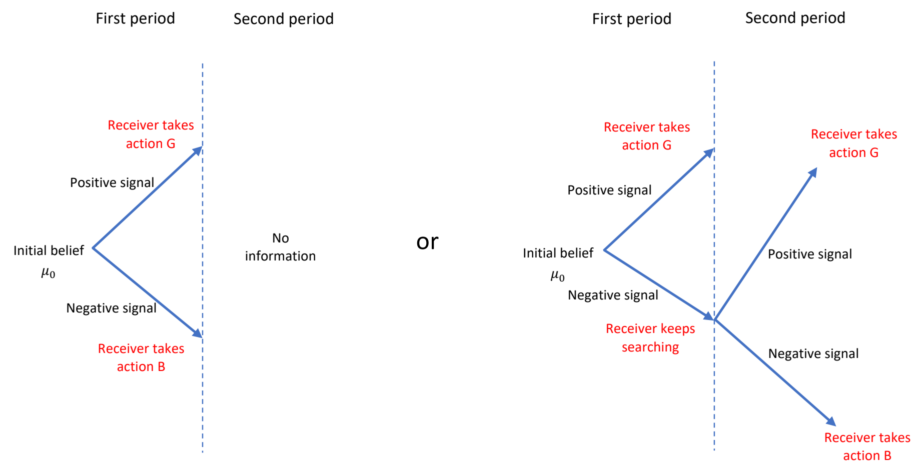

Figure 3: Belief Evolution under a One-shot Signal (Left) or a Pair of One-shot Signals (Right)
Left Figure: Sender Only Persuades in One Period; Right Figure: Sender Persuades in Both Periods

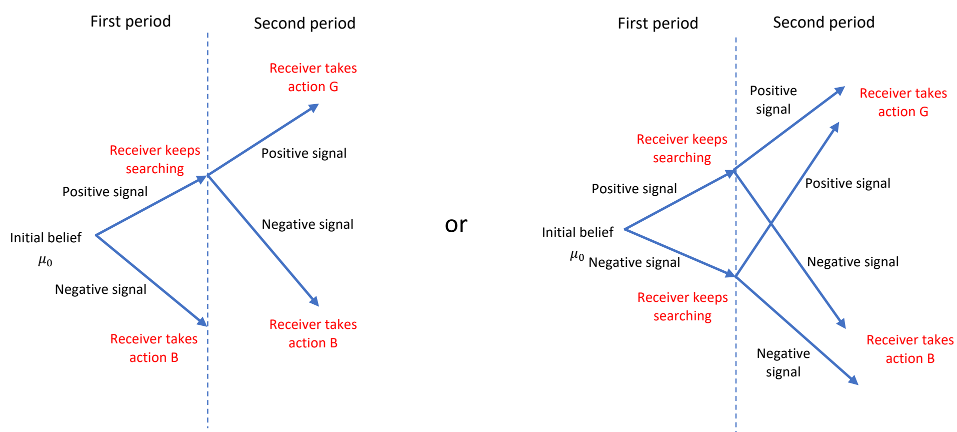

Figure 4: Belief Evolution under Iterative Signals
Left Figure: Receiver Keeps Searching Only after a Positive Signal; Right Figure: Receiver Searches in Both Periods

Compared to one-shot signals, iterative signals require a longer search time. To compensate the receiver for the higher expected search costs, the sender needs to give the receiver a higher benefit from searching to persuade him to search, which hurts the sender's payoff. The following result shows that the sender always prefers a one-shot signal to a pair of iterative signals in equilibrium. So, we limit our attention to the optimal one-shot signals in all subsequent analyses.

Proposition 1. *For any pair of feasible iterative signals, there exists a one-shot signal that gives the sender a strictly higher payoff.*

Outline of the proof.

1. Characterize the optimal one-period strategy

∘ Transform the sender's constrained non-linear programming with multiple constrained variables into an equivalent program with one constrained variable (λ_0).

Method: Replace $\bar{\mu_0}, \underline{\mu}_0$ by λ_0, and derive a lower and upper bound for λ_0 using (IR_0) and (F_0). Those bounds are necessary conditions of the original constraints.

Construct two feasible information structures that obtain the lower and upper bound of λ_0. It implies that the necessary conditions are also sufficient.

∘ Solve the transformed program using standard techniques.

2. For any feasible iterative signals, construct a feasible one-period signal that gives the sender a strictly higher payoff.

□

This proposition greatly simplifies the analyses by reducing the signal spaces that may be optimal and allows us to formulate the sender's problem in a unified way. The intuition of this proposition is that if the receiver is not certain enough to take action G after seeing a positive signal in the first period, it is better for the sender to change to another signal structure where the receiver would take action G right after seeing the positive signal. That is, the game tree of Figure 4 is dominated by the game tree on the left of Figure 3.[8] The sender takes into account the following constraints when designing the optimal one-shot signal of the first period:
(1) participation constraint:

$$\begin{aligned}&\lambda_0[\bar{\mu_0} v_g + (1-\bar{\mu_0})v_b] + (1-\lambda_0)\mathbb{E}[\text{receiver surplus at } t=1 | \text{search at } t=1]\\ =&\lambda_0(\bar{\mu_0}+v_b) + (1-\lambda_0)[\lambda_1(\bar{\mu_1}+v_b)-c] \geq c \qquad (IR_0)\end{aligned}$$

(2) feasibility constraint:

$$\lambda_0\bar{\mu_0} + (1-\lambda_0)\underline{\mu}_0 = \mu_0 \qquad (F_0)$$

If the sender provides information in both periods, her problem is:[9]

$$\begin{aligned}&\max_{\lambda_0,\bar{\mu_0},\mu_1,\lambda_1,\bar{\mu_1}} -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\lambda_1)+p\lambda_1\right] \qquad (P_2)\\ &\text{s.t. } (IR_0), (F_0), (\lambda_1,\bar{\mu}_1) \text{ solves } (P_1)\end{aligned}$$

The constraint that $(\lambda_1, \bar{\mu}_1)$ solves (P_1) implies that $(\lambda_1, \bar{\mu}_1)$ needs to satisfy the participation constraint, (IR_1), and feasibility constraint, (F_1), of (P_1). We use $(\lambda_0, \bar{\mu_0}, \mu_1, \lambda_1, \bar{\mu_1})$, the probability of a positive signal in each period, the belief after observing a positive signal in each period, and the

[8] We thank one of the anonymous reviewers for suggesting this intuition.

[9] To simplify the notation, we omit the following common constraints in the main text in all of the programs: $\bar{\mu_t} \in [-v_b, 1], \underline{\mu_t} \in [0, \mu_t), \lambda_t \in [0,1], \underline{\mu}_0 = \mu_1$. The first constraint is required by one-shot signals. The second and third constraints ensure the information structure is well-defined. The last equality comes from the fact that the belief at the beginning of the second period, μ_1, is the belief after observing a negative signal in the first period, $\underline{\mu}_0$, under one-shot signals.

initial belief in the second period, to represent the sender's strategy. We discuss some benchmark problems in the next section, and then analyze the solution to our problem in section 4.

3 Some Benchmarks

Before solving our model in the next section, we will work out several benchmarks to illustrate the added value of our model.

3.1 Standard Single-period Bayesian Persuasion Model

The first benchmark is the standard single-period Bayesian persuasion model in which there is no cost of providing or searching for information, and the receiver is forced to search. The sender does not need to consider the receiver's participation constraint. The sender can choose any signal to persuade the receiver to take action G, as long as the posterior belief is a mean-preserving spread of the prior belief (satisfies the feasibility constraint).

Standard Concavification Method

The above information design problem can be solved by the concavification method introduced by Aumann and Maschler (1995) and Kamenica and Gentzkow (2011). Figure 5 illustrates the solution method. The orange line is the sender's payoff without any information. Denote it by $\hat{v}(\mu)$. Given the belief μ, the receiver's payoff from taking action G is $\mu v_g + (1-\mu)v_b$ while his payoff from taking action B is 0. The receiver will take action G if and only if the belief is high enough, $\mu v_g + (1-\mu)v_b \geq 0 \Leftrightarrow \mu \geq -v_b$. So, the sender's payoff is 0 (p) if the belief is lower (higher) than $-v_b$. The sender will not benefit from providing partial information if the belief is higher than $-v_b$. In contrast, she can be better off providing information if the belief is lower. For example, by designing information such that the receiver will get a positive signal with probability $\frac{\mu_0}{-v_b}$ and raise his belief to $-v_b$, and a negative signal with probability $1-\frac{\mu_0}{-v_b}$ and lower his belief to 0, the sender can obtain a payoff of $\frac{\mu_0}{-v_b}p + (1-\frac{\mu_0}{-v_b}) \cdot 0 = \frac{\mu_0}{-v_b}p$. More generally, there is a one-to-one mapping from the posterior belief to the sender's payoff. So, her possible payoff with information design is the weighted sum of her payoff under different posterior beliefs, or formally, the convex hull of the graph of $\hat{v}(\mu)$, $co(\hat{v}(\mu))$. The highest payoff she can obtain is thus the concave closure of $\hat{v}(\mu)$, $\Pi(\mu) = sup\{u : (\mu, u) \in co(\hat{v}(\mu))\}$ (dashed blue line).

Why Concavification Does Not Work When the Receiver Searches Voluntarily

The above concavification method does not work when the receiver searches voluntarily because the sender's payoff with information design is no longer a weighted sum of her payoff under different posterior beliefs. It also depends on whether the receiver searches for information. For instance, consider the information structure discussed above. If the receiver receives a positive signal, he will take action G and obtain an expected payoff of $-v_b v_g + (1+v_b)v_b = 0$. If he observes a negative signal, he will take action B and obtain zero payoff. His expected payoff by searching for information is thus $0 - c < 0$. So, the receiver will take action B without searching. The sender's

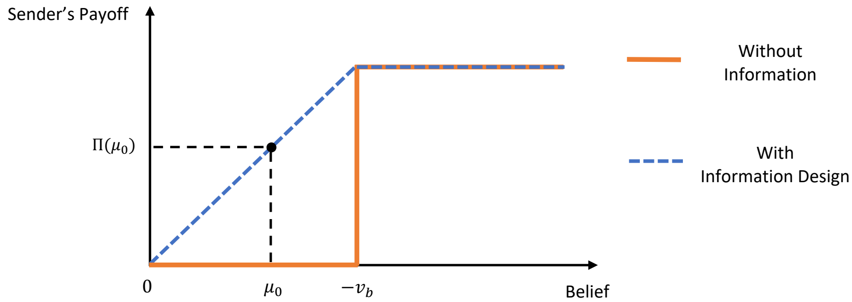

Figure 5: The Sender's Payoff under Standard Bayesian Persuasion

payoff would be 0 rather than $\Pi(\mu_0)$. This example illustrates that the sender's possible payoff is no longer the convex hull of the graph of $\hat{v}(\mu)$. Therefore, she may not achieve a payoff of the concave closure of $\hat{v}(\mu)$.

Consumer voluntary search is ubiquitous in marketing, as firms usually can only attract rather than force consumers to visit their website, click on their ad, or visit their store. The example above shows that the concavification method is not compatible with voluntary search, and that the optimal strategy and payoff may be qualitatively different. For example, the sender will always add some noise to the positive signal in this benchmark. In contrast, she may provide a precise positive signal ($\bar{\mu} = 1$) in our model.

How Our Method Solves the Problem

This paper develops an alternative method to solve this problem and, more broadly, other two-player games in which voluntary participation is voluntary. It directly incorporates the receiver's participation constraints in the setup. In this simple example, the sender's constrained non-linear programming is:

$$\max_{\lambda, \bar{\mu}} \; p\lambda \text{ s.t. } \lambda(\bar{\mu} + v_b) \geq c, \lambda\bar{\mu} + (1-\lambda)\underline{\mu} = \mu, \lambda \in [0,1], \underline{\mu} \in [0, \mu)$$

There are two constrained variables, λ and $\bar{\mu}$. The challenges are that some constraints are non-linear, and both constrained variables appear jointly in some constraints. We solve this problem by transforming the above program into an equivalent program with one constrained variable and linear constraints (when $\mu_0 \geq c/v_g$):

$$\max_{\lambda} \; p\lambda \text{ s.t. } \lambda \in \left[\frac{c}{v_g}, \frac{\mu_0 - c}{-v_b}\right]$$

It then becomes immediate that the optimal λ is $\frac{\mu_0 - c}{-v_b}$. We can then find the corresponding $\bar{\mu}$ to pin down the optimal information structure. Due to the players ' intertemporal tradeoff, the dynamic problem has more constrained variables, more complicated constraints, and a much larger signal space. But we can extend the method in this section to that case. Please refer to the proof outlines of Proposition 1, 5, and 6 for more details.

3.2 Constant Information Provision Costs

In the second benchmark, we consider the same model as ours, except that the sender's information provision cost does not depend on the specific information structure. If the sender provides information, she incurs a constant cost of k in that period. The following result characterizes the optimal information structure. It shows that the sender will never smooth information over two periods.

Proposition 2. *The sender does not provide information if the prior is low,* $\mu_0 < \max\{c/v_g, c - v_b k/p\}$*, and provides information in one period if* $\mu_0 \geq \max\{c/v_g, c - v_b k/p\}$*. In the latter case, the optimal probability of a positive signal and the optimal belief after observing a positive signal,* $(\lambda_0^*, \bar{\mu}_0^*)$*, is* $(\frac{\mu_0 - c}{-v_b}, \frac{-v_b \mu_0}{\mu_0 - c})$.

This proposition shows that the convexity of the information provision cost is important for information smoothing. Intuitively, the sender can save some information provision costs by smoothing the information over two periods if the cost is convex. If the cost is constant instead, the sender can always combine the signals in both periods into a single signal and only incurs the information provision cost once. That way, she will be strictly better off. It does not mean that the information provision cost is always convex. A different cost function may be another reason the sender sometimes only communicates with the receiver once. However, it implies that the convexity of the information provision cost plays a critical role when the sender persuades the receiver gradually.[10]

3.3 Alternative Non-Bayesian Persuasion Model

Bayesian Persuasion models allow the sender to choose flexibly from a large signal space. An alternative communication model is information disclosure, where the sender can only disclose information completely or nothing. Figure 6 illustrates its optimal solution. Again, the orange line is the sender's payoff without any information. The receiver will take action G if and only if the belief is high enough. The sender will not benefit from providing information if the belief is higher than $-v_b$. In contrast, she can be better off providing information if the belief is lower. Because the sender can not provide partial information, the receiver will know the exact state after disclosure. With probability μ_0, the receiver knows that the state is g and takes action G; with probability $1 - \mu_0$, he knows that the state is b and takes action B. The sender's payoff is thus $\mu_0 \cdot p$ with disclosure. The green line illustrates the sender's payoff with optimal disclosure.

Comparing Figure 5 with Figure 6, one can see that the ability to design flexible information is valuable to the sender. Disclosure is, in general, sub-optimal for her if she has access to a larger signal space. In the real world, firms often have at least some freedom in choosing how much and what kinds of information to provide to consumers. This benchmark highlights the value of a Bayesian persuasion model, which can offer us more insights into the information design problem.

[10] We thank one of the anonymous reviewers for suggesting this benchmark.

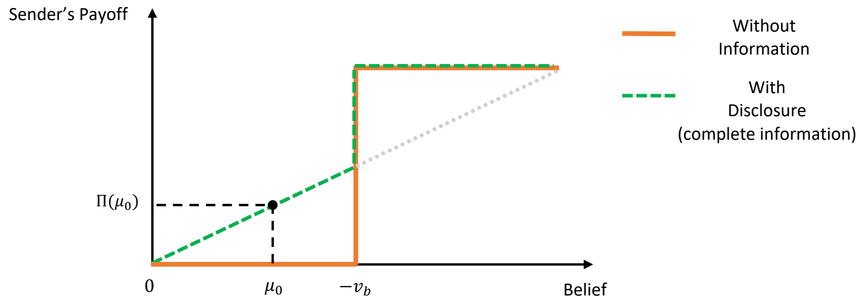

Figure 6: The Sender's Payoff under Information Disclosure

3.4 Dynamic Commitment

Under many circumstances, the assumption that the sender can generate credible signals within each period but does not have dynamic commitment power is reasonable. It is hard for the sender to commit to the entire information structure across all periods and convince the receiver that she will stick to it when deviating to a different information structure during intermediate periods is profitable. However, factors such as reputation can give the sender stronger commitment power. Here, we study the implications of dynamic commitment power. The sender chooses and commits to the entire information structure to maximize the expected surplus.

$$\max_{\lambda_0,\bar{\mu_0},\mu_1,\lambda_1,\bar{\mu_1}} -K(\lambda_0)+p\lambda_0+(1-\lambda_0)\left[-K(\lambda_1)+p\lambda_1\right] \tag{P_{dc}}$$
$$\text{s.t. } (IR_0), (F_0), (IR_1), (F_1)$$

Compared to (P_2), the sender faces one fewer constraint if she has dynamic commitment power: the second-period information does not need to maximize the sender's second-period payoff. Consequently, the sender can access a larger signal space and will always be (weakly) better off. Since dynamic commitment does not matter if the sender only provides information once, we look at the case in which the sender provides information in both periods.

Proposition 3. *Suppose the search cost is high, $v_g\lambda_1^{**} < c < \widehat{c}$, and $\mu_0 > c(2v_g - c)/(v_g)^2$. The sender provides information in both periods regardless of the dynamic commitment power. If the sender has dynamic commitment power, her payoff is strictly higher, and the receiver gets a strictly positive surplus in the second period. The benefit of dynamic commitment power for the sender vanishes as the search cost approaches zero.*

When the search cost is high, we will show in Proposition 6 that the sender perfectly smooths the information (the same λ and $\bar{\mu}$ in each period) if she does not have dynamic commitment power. If she instead has dynamic commitment power, she will commit to giving the receiver a higher benefit from searching in the second period. As a result, the receiver will search even if the amount of persuasion is lower in the first period. Though it hurts the sender's payoff in the second period, it increases her payoff in the first period by reducing the information provision cost. The overall effect is strictly positive. So, the optimal information provision will not be perfectly smooth. The above finding relates to the durable good pricing results (e.g., Coase 1972), though the underlying mechanisms differ. Without dynamic commitment power, the monopolist tends to reduce the price as time passes, which reduces profit because a rational receiver will strategically wait. Here, the ability to commit to more favorable (higher expected surplus) information in the future benefits the sender.

In contrast, when the search cost approaches zero, the difference between the sender surplus with and without dynamic commitment power approaches zero. Intuitively, the sender can commit to giving the receiver a higher benefit from searching in the second period if she has dynamic power. It benefits her by relaxing the receiver's participation constraint in the first period. However, the participation constraint is already very loose when the search cost is low. Thus, the benefit of dynamic commitment power approaches zero.

The sender needs to have more strategic considerations if she does not have dynamic commitment power. This benchmark shows that this additional strategic consideration is vital in the presence of non-trivial friction (search cost).

4 Optimal Strategies

From the discussion in Section 2.3.1, the receiver will search (take action S) whenever the sender provides information. Otherwise, the sender can save the information provision cost and be better off by not providing any information. The receiver will take action G immediately upon receiving a positive signal and action B if the sender does not provide information. Therefore, we only need to characterize the sender's strategy, which implies the receiver's strategy.

4.1 A Relaxed Problem

We first consider a relaxed problem of the original problem. In the relaxed problem (P_r), The sender maximizes the objective function of the original problem without taking into account the participation constraints and the feasibility constraints. We will use the solution to the relaxed problem throughout the subsequent analyses.

$$\max_{\lambda_0,\lambda_1} -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\lambda_1) + p\lambda_1\right] \tag{P_r}$$

Lemma 1. *The solution to the relaxed problem (P_r), $(\lambda_0^{**}, \lambda_1^{**})$, does not depend on the search cost c and $\lambda_0^{**} < \lambda_1^{**}$. The sender's payoff in the relaxed problem is strictly positive.*

The above payoff is the highest possible payoff the sender can obtain in equilibrium. When the prior belief is high enough, the sender obtains that payoff by setting the probability of a positive signal to λ_t^{**}. Since the sender can smooth the information provision at the beginning of the first period while only has one chance of providing information in the second period, she will choose $\lambda_0^{**} < \lambda_1^{**}$. When the prior is lower and a positive signal occurs with probability λ_0^{**} and λ_1^{**} in each period, the sender needs to provide a very noisy signal (low $\bar{\mu_t}$) due to feasibility constraints. As a result, the receiver will be quite uncertain about the state even after observing a positive signal. So, he will choose not to search for information. This friction restricts the communication between the players and distorts the optimal strategy away from the strategy of the relaxed problem. For the problem to be non-degenerate, we concentrate on the case in which the prior is not too high throughout the paper. As a result, the optimal strategy is different from from the strategy of the relaxed problem.

4.2 Optimal Strategy in the Second Period

When the belief at the beginning of the second period is too low, the receiver will not search, given any feasible signals. Thus, the sender does not provide information to minimize the cost. When the belief at the beginning of the second period is higher, and the search cost is not too high, the sender provides information and obtains a positive payoff. The following proposition summarizes the optimal information structure.

Proposition 4. *In the second period, the sender does not provide information when $\mu_1 < \mu_{0,1} := c/v_g$. When $\mu_1 \geq \mu_{0,1}$, the optimal probability of a positive signal and the optimal belief after observing a positive signal, $(\lambda_1^*, \bar{\mu}_1^*)$, depend on the search cost c:*

1. *If $c \geq v_g\lambda_1^{**}$, there exists a unique $\widehat{c} \in (v_g\lambda_1^{**}, v_g\mu_1]$ such that the sender does not provide information if $c > \widehat{c}$ and $(\lambda_1^*, \bar{\mu}_1^*) = (c/v_g, 1)$ if $c < \widehat{c}$. The receiver gets zero surplus.*
2. *If $c \in [\mu_1 + v_b\lambda_1^{**}, v_g\lambda_1^{**})$, $(\lambda_1^*, \bar{\mu}_1^*) = (\frac{\mu_1 - c}{-v_b}, \frac{-v_b\mu_1}{\mu_1 - c})$. The receiver gets zero surplus.*
3. *If $c < \min\{\mu_1 + v_b\lambda_1^{**}, v_g\lambda_1^{**}\}$, $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \min\{\frac{\mu_1}{\lambda_1^{**}}, 1\})$. The receiver gets positive surplus.*

When the search cost is too high, the sender has to incur a very high cost to persuade the receiver to search. Even if it is feasible for the sender to induce the receiver to search, it is so costly that the sender's expected payoff is negative. So, the sender chooses not to provide information, and the receiver does not search.

When the search cost is high but not too high, the sender will provide just enough information such that the receiver searches. Since the sender needs to give the receiver a high benefit from searching to persuade him to search, in equilibrium, the receiver becomes certain ($\bar{\mu_1} = 1$) that the state is g after observing a positive signal. Suppose, instead, a positive signal does not fully reveal the state ($\bar{\mu_1} < 1$). In that case, its arrival rate will need to be higher to persuade the receiver to search. Since the marginal cost of increasing the probability of a positive signal exceeds the marginal benefit, the sender's payoff decreases. The sender trades off the frequency of positive signal for precision.

When the search cost is medium, the receiver's participation constraint is easier to satisfy. Since the marginal benefit of increasing the probability of a positive signal exceeds the marginal cost, in equilibrium, the sender trades off the precision of a positive signal for frequency. The receiver is still uncertain about the state after observing a positive signal, but the belief is high enough that the receiver searches.

When the search cost is low, the information friction does not distort the information structure from the solution to the relaxed problem, and the receiver gets a strictly positive surplus.

Figure 7 and Figure 8 illustrate the optimal strategy and receiver surplus as a function of the search cost. For low search cost in both figures, the sender can achieve the solution to the relaxed problem and obtain the highest possible payoff because the communication friction is low. As the search cost increases, the sender needs to give the receiver a higher benefit from searching to satisfy his participation constraint. In Figure 7, the sender cannot increase λ_1 or $\bar{\mu}_1$ without lowering the other one due to the low initial belief μ_1 ($\underline{\mu}_1$ is already 0 and cannot be further reduced, so $\lambda \cdot \bar{\mu}$ is fixed). When she switches from the low search cost region to the medium search cost region, she reduces the likelihood of a positive signal and increases its precision. It shows that the receiver surplus is higher when the positive signal is more precise but less likely, fixing $\lambda \cdot \bar{\mu}$. In Figure 8, the sender can increase λ_1 without lowering $\bar{\mu}_1$ due to the high initial belief μ_1 ($\underline{\mu}_1$ is positive, so she can increase $\lambda \cdot \bar{\mu}$ by reducing $\underline{\mu}_1$). When she switches from the low search cost region to the high search cost region, she increases the likelihood of a positive signal to satisfy the receiver's participation constraint.[11]

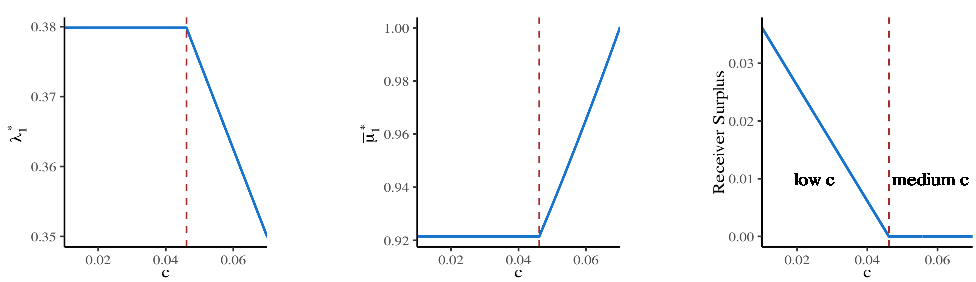

Figure 7: The sender's optimal strategy and receiver surplus when $c = 0.01, \mu_1 = 0.35, p = 0.8, v_g = 0.2, v_b = -0.8, K(\lambda) = \frac{1}{2}\frac{\lambda^2}{1-\lambda}$.

4.3 Optimal Strategy for the Entire Game

When the prior is too low, any feasible signal the sender can generate is not attractive enough for the receiver to search. Thus, it is impossible to communicate between the sender and the receiver. When the prior is higher, and the search cost is not too high, the sender provides information and obtains a positive payoff.

[11] Note that the medium search cost region is empty given the parameter values in this figure.

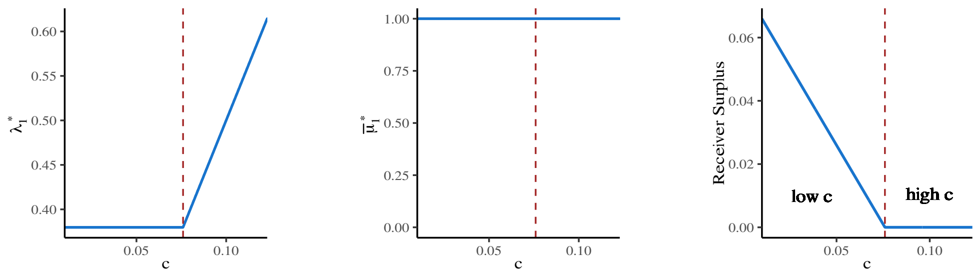

Figure 8: The sender's optimal strategy and receiver surplus when $c = 0.01, \mu_1 = 0.6, p = 0.8, v_g = 0.2, v_b = -0.8, K(\lambda) = \frac{1}{2}\frac{\lambda^2}{1-\lambda}$.

Proposition 5. *Suppose the search cost is not too high, $c < \widehat{c}$. There exists $\mu_{1,2} \geq c(2v_g - c)/(v_g)^2$ such that the sender does not provide information if the prior is low, $\mu_0 < \mu_{0,1}$, provides information in one period if $\mu_0 \in [\mu_{0,1}, \mu_{1,2})$, and provides information in both periods if $\mu_0 > \mu_{1,2}$. Suppose the sender provides information in both periods. A positive signal fully reveals the state being g when the search cost is high and partially reveals the state when the search cost is low. The receiver gets zero total surplus.*

As we discussed in the previous section, the widely-used concavification approach cannot be used to solve this kind of games because the receiver's participation is strategic. We instead develop a constrained non-linear programming method to solve the sender's information design problem.

Outline of the proof.

1. $v_g\lambda_1^{**} \leq c < \widehat{c}$

   (a) Characterize the optimal two-period strategy

   ○ Proposition 4 implies that the optimal amount of persuasion in the second period, λ_1^*, is c/v_g. Replace λ_1 by c/v_g in the objective function of the sender's program.

   ○ Transform the sender's program with multiple constrained variables into an equivalent program with one constrained variable (λ_0). (Similar method as characterizing the optimal one-period strategy in the proof of Proposition 1.)

   ○ Solve the transformed program using standard techniques.

   (b) Compare the sender's payoff from the optimal one-period and two-period strategies. Show that the sender always provides information in both periods as long as feasible (the prior belief is high enough).

2. $c < v_g\lambda_1^{**}$

   (a) Characterize the optimal two-period strategy

∘ Transform the sender's program with multiple constrained variables into two constrained programs.

We need to transform the original program into two programs because the receiver's participation constraint in the first period depends on his expected payoff in the second period. Denote the corresponding strategy as the S_+ (S_0) strategy and the corresponding sender's program as (P_{2S_+}) ((P_{2S_0})) if the sender gives the receiver a strictly positive surplus (zero surplus) in the second period.

∘ Obtain the solution to each program

i. ∘ Transform (P_{2S_+}) into an equivalent program with one constrained variable (λ_0). (similar method as before)

∘ Solve the transformed program using standard techniques.

ii. ∘ Transform (P_{2S_0}) into an equivalent program with one constrained variable (μ_1). We cannot use a similar method as before to transform (P_{2S_0}) into an equivalent program with one constrained variable (λ_0) because μ_1 also appears in the objective function. Instead, we transform (P_{2S_0}) into an equivalent program with two constraints specifying the lower and upper bound of λ_0 and μ_1. We then prove λ_0 is binding at the upper bound. It allows us to replace λ_0 by its upper bound in the objective function and omit it in the constraints.

∘ Solve the transformed program using standard techniques.

(b) Compare the sender's payoff from the optimal one-period and two-period strategies.

∘ Show that the sender's payoff from the optimal S_+ strategy is always higher than her payoff from the optimal one-period strategy.

∘ Prove a single-crossing result: if the sender's payoff from the optimal S_0 strategy is higher than her payoff from the optimal one-period strategy at a given prior, then her payoff from the optimal S_0 strategy is higher than her payoff from the optimal one-period strategy at any higher prior.

It is proved by showing that the sender's payoff from the optimal S_0 strategy increases in μ_0 at a higher rate than her payoff from the optimal one-period strategy.

∘ The above two results imply that the sender will always provide information in both periods if and only if the prior is higher than a threshold.

□

Since the information provision cost is convex in the probability of a positive signal, the sender has an incentive to smooth the information provision over two periods. When the prior is low, it is not feasible for the sender to provide enough persuasion in both periods so that the receiver will search whenever a positive signal has not arrived. As the prior increases, it becomes feasible for the sender to smooth the information provision. If the sender finds it optimal to provide information in both periods at a given prior, she also prefers to smooth the information for any higher prior.

The receiver obtains the highest possible surplus conditional on observing a positive signal and making the purchase. When it is highly costly for the receiver to search, the optimal information structure fully convinces the receiver that the state is g when a positive signal arrives. Without providing this type of information, the sender cannot persuade the receiver to search. In contrast, when it is less costly for the receiver to search, the optimal information structure adds some noise to a positive signal. In equilibrium, the receiver is not sure that the state is g after observing a positive signal. The state may be b after the receiver takes action G. However, the likelihood of state g after a positive signal is high enough to persuade the receiver to search. By adding some noise to a positive signal, the sender can generate more frequent positive signals and increase her payoff without violating the feasibility constraint.

When the sender provides information in both periods, she can always extract surplus from the receiver if the receiver gets a strictly positive surplus. If the likelihood of a positive signal is lower than its unconstrained optimum, the sender can increase her payoff by increasing the probability of a positive signal and decreasing the belief after observing a positive signal. If the likelihood of a positive signal is higher than its unconstrained optimum, the sender can increase the payoff by reducing the probability of a positive signal and increasing the belief after observing a positive signal. This implies that the receiver gets zero surplus in equilibrium.

4.4 Comparative Statics

When the sender provides information in only one period, the optimal strategy has a closed-form solution and is easy to analyze. Here, we discuss the comparative statics when the sender provides information in both periods.

4.4.1 Comparative Statics With Regard to the Prior Belief

When the search cost is high, the prior determines whether the sender smooths information but does not affect the information structure, conditional on the sender providing information. When the search cost is low, the prior affects the information structure monotonically. When the search cost is intermediate, the sender may switch from the S_0 strategy to the S_+ strategy as the prior increases. There can be a discrete jump in the optimal information structure. We leave the analysis of this case to the online appendix.

Proposition 6. *Suppose the sender provides information in both periods. When the search cost is high, $v_g\lambda_1^{**} \leq c < \widehat{c}$, positive signal fully reveals the state. Neither the probability of a positive signal, λ_t^*, nor the sender's payoff depends on the prior, $(\lambda_t^*, \bar{\mu}_t{}^*) = (c/v_g, 1)$. When the search cost is low, $c \leq \tilde{c} := v_g K'^{-1}\left[K(\lambda_1^{**})/\lambda_1^{**}\right]$, the probability of a positive signal, λ_t^*, is continuous and increases in the prior. The belief after observing a positive signal, $\bar{\mu}_t^*$, is continuous and decreases in the prior. The sender's payoff strictly increases in the prior.*

Outline of the proof.

The proof of Proposition 5 has fully characterized the optimal strategy when the search cost is high. To fully characterize the optimal strategy when the search cost is low, we only need to select the global solution by comparing the local solutions to (P_{2S_+}) and (P_{2S_0}) when both are feasible.

1. High prior belief, $\mu_0 \geq \mu_0'$ (μ_0' is defined in the appendix)

   ∘ Show that the optimal S_+ strategy also satisfies the constraints in (P_{2S_0}). Thus, the sender's payoff from the optimal S_0 strategy is higher than her payoff from the optimal S_+ strategy.

2. Low prior belief

   ∘ Show that the sender's payoff from the optimal S_+ strategy increases in μ_0 at a higher rate than her payoff from the optimal S_0 strategy.

   According to the high prior case, the sender's payoff from the optimal S_+ strategy is lower than her payoff from the optimal S_0 strategy when $\mu_0 = \mu_0'$. Hence, her payoff from the optimal S_+ strategy is lower than that from the optimal S_0 strategy for any $\mu_0 < \mu_0'$.

∘ Show that S_0 strategy is feasible whenever S_+ strategy is feasible. Therefore, the sender always uses the S_0 strategy when she provides information in both periods.

∘ Derive the comparative statics by first-order conditions and the implicit function theorem. □

The optimal information is perfectly smooth (the same λ and $\bar{\mu}$ in each period) when the search cost is high and the sender provides information in both periods. Since the receiver's participation constraint is strong, a positive signal fully reveals the state is g to give him enough benefit from searching. Because it is very costly for the receiver to acquire information, the minimal amount of persuasion to persuade him to search is high. The marginal cost of increasing the amount of persuasion exceeds the marginal benefit. Even if the prior increases and it is feasible for the sender to increase the amount of persuasion, she will prefer not to do so. Hence, conditional on the sender providing information, the information structure does not depend on the prior. The positive signal has the same likelihood and precision in each period. Figure 9 illustrates the sender's optimal strategy numerically when the search cost is high.[12] As we can see, the optimal information structure is perfectly smooth, and a positive signal always fully reveals the state ($\bar{\mu}_0 = \bar{\mu}_1 = 1$).

[12] The choice of the specific parameter values does not affect the qualitative property of the optimal strategy (i.e., the shape of the figure). What matters is the relative value. The search cost is low if $c \leq v_g \lambda_0^{**}$, intermediate if $v_g \lambda_0^{**} < c < v_g \lambda_1^{**}$, and high if $v_g \lambda_1^{**} \leq c < \widehat{c}$.

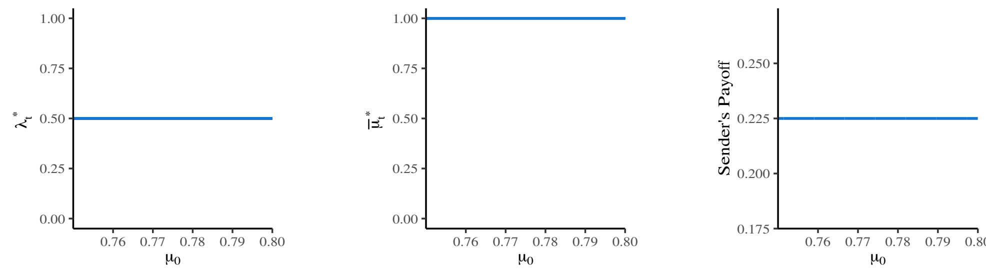

Figure 9: The sender's optimal strategy and payoff when $c = 0.1, p = 0.8, v_g = 0.2, v_b = -0.8, K(\lambda) = \frac{1}{2}\frac{\lambda^2}{1-\lambda}$.

When the search cost is low, and the sender provides information in both periods, she chooses between S_+ and S_0 strategies. Under the S_+ strategy, the receiver observes less frequent positive signals in the first period and more frequent positive signals in the second period. On average, he spends a longer time searching. Consequently, the sender has to give the receiver a higher benefit from searching to compensate for the higher expected total search cost, which reduces the sender's surplus. Therefore, the sender always chooses the S_0 strategy in equilibrium, and the optimal strategy is continuous in the prior. The likelihood of a positive signal is lower than its unconstrained optimum in both periods. More frequent positive signals are feasible when the prior is higher. Even if the receiver becomes less sure about the state being good after observing a positive signal, he will still search as long as the likelihood of receiving a positive signal and earning a strictly positive surplus increases. In equilibrium, the sender trades off the precision of a positive signal for frequency as the prior increases. The consumer spends less time searching for information because he is more likely to receive a positive signal and make a decision in the first period.

Figure 10 illustrates the sender's optimal strategy numerically when the search cost is low.[13] We present the optimal strategy and the sender's payoffs from the optimal one-period, S_0, and S_+ strategies. The sender always prefers the S_0 strategy when she provides information in both periods. As illustrated, the probabilities of a positive signal at both periods, λ_0^* and λ_1^*, are continuous and increase in μ_0. The beliefs after observing a positive signal in each period, $\bar{\mu}_0^*$ and $\bar{\mu}_1^*$, are continuous and decrease in μ_0. The positive signal has different likelihood and precision in each period. When the prior is lower than the intercept of the brown line, the sender prefers providing information in only one period to providing information in both periods.

[13] The domain of the prior is $[c(2v_g - c)/(v_g)^2, \min\{\widehat{\mu}_0, p\}]$. When $\mu_0 < c(2v_g - c)/(v_g)^2$, the sender provides information in at most one period. When $\mu_0 \geq \widehat{\mu}_0 := 2c - v_b\lambda_1^{**} - [c + (1-\lambda_1^{**})v_b]\lambda_0^{**}$, the sender achieves the solution to the relaxed problem.

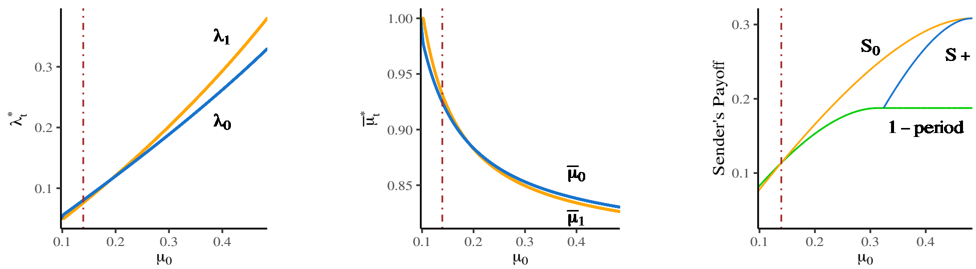

Figure 10: The sender's optimal strategy and payoff when $c = 0.01, p = 0.8, v_g = 0.2, v_b = -0.8, K(\lambda) = \frac{1}{2}\frac{\lambda^2}{1-\lambda}$.

4.4.2 Comparative Statics With Regard to the Sender's Costs

Providing the same amount of persuasion may impose different costs on the sender. To study the impact of the sender's information provision costs on the optimal strategy, we rewrite the sender's cost function as $K(\lambda) = \eta\tilde{K}(\lambda)$, where $\tilde{K}(1/2) = 1/2$ for identification. It is more costly for the sender to provide information when η is larger. The following proposition summarizes the comparative statics of the optimal strategy about η.

Proposition 7. *Suppose the sender provides information in both periods. Her payoff strictly decreases in the relative cost of information provision, η. When the search cost is low, $c \leq \tilde{c}$, the amount of persuasion decreases in the first period and increases in the second period, as η increases. When the search cost is high, $c \geq v_g\lambda_1^{**}$, the optimal strategy of the sender does not depend on η.*

When the search cost is high, it is very costly for the receiver to search. The sender needs to incur a high cost to persuade the receiver to search. As a result, the marginal cost of increasing the amount of persuasion exceeds the marginal benefit. So, the sender provides the minimum amount of persuasion for the receiver to search, which does not depend on the sender's cost. Hence, the optimal information structure does not depend on the relative cost of information provision.

When the search cost is low, as the relative cost of information provision increases, the marginal cost of increasing the amount of persuasion increases, while the marginal benefit remains the same. Because the sender definitely incurs the information provision cost in the first period, she reduces the amount of persuasion in the first period. This allows her to increase the amount of persuasion in the second period when the information provision cost is not always incurred and the likelihood of a positive signal is lower than its unconstrained optimum. The consumer spends more time searching for information because he is less likely to receive a positive signal and make a decision in the first period. Figure 11 and 12 illustrate the sender's optimal strategy when the search cost is high and low, respectively.

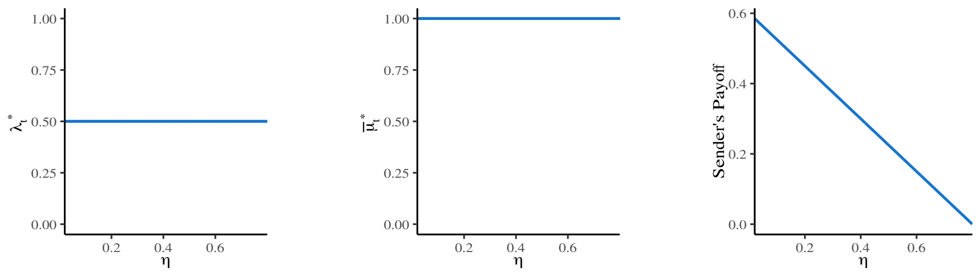

Figure 11: The sender's optimal strategy and payoff when $c = 0.1, \mu_0 = 0.78, p = 0.8, v_g = 0.2, v_b = -0.8, K(\lambda) = \eta\frac{\lambda^2}{1-\lambda}$.

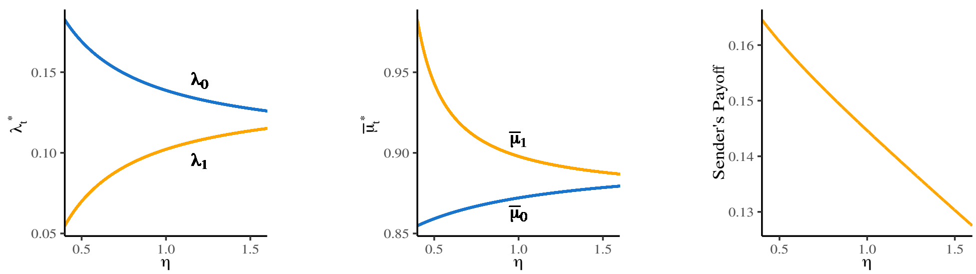

Figure 12: The sender's optimal strategy and payoff when $c = 0.01, \mu_0 = 0.2, p = 0.8, v_g = 0.2, v_b = -0.8, K(\lambda) = \eta\frac{\lambda^2}{1-\lambda}$.

5 The Efficient Information Structure

In the previous sections, the sender designs the information structure to maximize the expected payoff. This section characterizes the efficient strategy when a social planner designs the information structure to maximize total welfare. We then investigate the information distortion caused by not taking into account receiver surplus. For tractability reasons, we use a special form of the payoff function in this section, $v_g = 1 - p$ and $v_b = -p$.

5.1 Efficient Strategy in the Second Period

As discussed in the previous section, the sender does not provide information if the belief at the beginning of the second period is too low, $\mu_1 < \mu_{0,1}$. So, we concentrate on the case in which $\mu_1 \geq \mu_{0,1}$. In the second period, the social planner's problem is:

$$\max_{\lambda_1, \bar{\mu_1}} -K(\lambda_1) + p\lambda_1 + \lambda_1(\bar{\mu_1} - p) - c \tag{E_1}$$
$$\text{s.t. } (IR_1), (F_1)$$

Below, we discuss the efficient information structure in the second period intuitively. The online appendix presents the formal characterization of the efficient strategy in the second period. When the search cost is high, the information provider must incur a high cost to persuade the receiver to search. The marginal cost of increasing the amount of persuasion exceeds the marginal benefit. Therefore, both the sender and the social planner choose the minimum amount of persuasion that can induce the receiver to search, and there is no information distortion. When the search cost is lower, it is easier to persuade the receiver to search. The marginal costs of increasing the amount of persuasion are the same for both the sender and the social planner, while the marginal benefit is lower for the sender. Therefore, the amount of persuasion the sender chooses is lower than the social planner does, except when both the search cost and the initial belief are low. In that case, the sender chooses a higher amount of persuasion than the social planner because it is not feasible for the latter to generate frequent positive signals.

5.2 Efficient Strategy for the Entire Game

As in the previous section, the social planner does not provide information if the prior is too low or the search cost is too high. When she provides information in only one period, the previous subsection characterizes the efficient strategy. When she provides information in both periods, her problem is:

$$\max_{\lambda_0, \bar{\mu_0}, \mu_1, \lambda_1, \bar{\mu_1}} -K(\lambda_0) + \lambda_0 \bar{\mu}_0 - c + (1 - \lambda_0)\left[-K(\lambda_1) + \lambda_1 - c\right] \tag{E_2}$$

$$\text{s.t. } (IR_0), (F_0), (\lambda_1, \bar{\mu}_1) \text{ solves } (E_1)$$

The following proposition compares the payoff-maximizing and efficient strategies when the sender provides information in both periods.

Proposition 8. *Suppose the sender provides information in both periods. When $c \geq v_g \lambda_1^{**}$, the payoff-maximizing strategy is efficient. When $c < v_g \lambda_1^{**}$, the sender, who maximizes her payoff, chooses a lower amount of persuasion in the first period and a higher one in the second period than does the social planner, who maximizes total welfare.*

There is no information distortion when the search cost is high, similar to the argument in the previous subsection. When the search cost is lower, the sender's information structure is no longer efficient because she does not internalize the receiver's welfare. Since the social planner benefits directly from reducing the receiver's search cost, she designs the information to speed up the search process. The receiver takes action G with a higher probability in the first period under the efficient information structure.

6 Concluding Remarks

Consumers frequently search for information before making decisions. Since their search and purchase decisions depend on the information environment, firms have a strong incentive to influence it. This paper endogenizes consumers' information environment from the firm's perspective under a general signal space. We examine the optimal information provision strategy of a sender and the optimal information acquisition strategy of a receiver when the sender sequentially persuades a receiver to take a particular action. The sender prefers that action regardless of the unknown state, while the receiver only wishes to take that action if the state is good. In our model, the sender incurs a cost to provide information, and the receiver incurs a cost to search. The receiver trades off the cost of searching and the benefit of obtaining more accurate information to make better decisions. The sender trades off the cost of information provision and the benefit of persuading the receiver to search and then take the sender's preferred action. We allow for gradual communication between the sender and the receiver. Consequently, the sender also makes the intertemporal trade-off of smoothing the information to reduce the persuasion cost.

In equilibrium, the sender uses one-shot signals that induce the receiver to immediately take the sender's preferred action upon observing a positive signal. The sender smooths information over multiple periods if and only if there is a high prior that the state is good. The sender extracts all the surplus from the receiver when she provides information in both periods, while she may leave some surplus for the receiver when she provides information in only one period. When the search cost for the receiver is high, the receiver is sure that the state is good when he takes the desired action. When the search cost is low, the optimal information structure does not fully reveal the state, which may be bad even though the receiver takes the desired action. We compare the payoff-maximizing information structure with the efficient information structure and find no information distortion when the search cost is high. There is information distortion when the search cost is lower.

There are some limitations to the current work. The implementation of a given signal depends on the institutional details of the specific problem. Further empirical work can complement the current paper by putting the theoretical results into practice. In addition, it will be interesting to study the optimal persuasion strategy when there is more than one sender. Such competition may lead the sender to provide more information and improve equilibrium efficiency. Moreover, the sender has complete control of the information structure in this paper. It will be interesting to consider the case where the sender can only partially control the information environment.

References

Aumann, R. J., Maschler, M. (1995). Repeated games with incomplete information. MIT press.

Ball, I. (2019). Dynamic information provision: Rewarding the past and guiding the future. Available at SSRN 3103127.

Ball, I., Espín-Sánchez, J. A. (2022). Experimental Persuasion.

Bergemann, D., Morris, S. (2019). Information design: A unified perspective. Journal of Economic Literature, 57(1), 44-95.

Berman, R., Zhao, H., Zhu, Y. (2022). Strategic Recommendation Algorithms: Overselling and Demarketing Information Designs. Available at SSRN 4301489.

Bizzotto, J., Rüdiger, J., Vigier, A. (2021). Dynamic persuasion with outside information. American Economic Journal: Microeconomics, 13(1), 179-94.

Branco, F., Sun, M., Villas-Boas, J. M. (2012). Optimal search for product information. Management Science, 58(11), 2037-2056.

Branco, F., Sun, M., Villas-Boas, J. M. (2016). Too much information? Information provision and search costs. Marketing Science, 35(4), 605-618.

Che, Y. K., Kim, K., Mierendorff, K. (2020). Keeping the listener engaged: a dynamic model of bayesian persuasion. arXiv preprint arXiv:2003.07338.

Chen, Y., Yao, S. (2017). Sequential search with refinement: Model and application with click-stream data. Management Science, 63(12), 4345-4365.

Coase, R. H. (1972). Durability and monopoly. The Journal of Law and Economics, 15(1), 143-149.

Degan, A., Li, M. (2021). Persuasion with costly precision. Economic Theory, 72(3), 869-908.

Dukes, A., Liu, L. (2016). Online shopping intermediaries: The strategic design of search environments. Management Science, 62(4), 1064-1077.

Ely, J. C. (2017). Beeps. American Economic Review, 107(1), 31-53.

Ely, J. C., Szydlowski, M. (2020). Moving the goalposts. Journal of Political Economy, 128(2), 468-506.

Gardete, P., Hunter, M. (2020). Guiding Consumers through Lemons and Peaches: An Analysis of the Effects of Search Design Activities.

Gentzkow, M., Kamenica, E. (2014). Costly persuasion. American Economic Review, 104(5), 457-62.

Guo, L. (2021). Endogenous Evaluation and Sequential Search. Marketing Science, 40(3), 413-427.

Hong, H., Shum, M. (2006). Using price distributions to estimate search costs. The RAND Journal of Economics, 37(2), 257-275.

Honka, E. (2014). Quantifying search and switching costs in the US auto insurance industry. The RAND Journal of Economics, 45(4), 847-884.

Honka, E., Chintagunta, P. (2017). Simultaneous or sequential? search strategies in the us auto insurance industry. Marketing Science, 36(1), 21-42.

Iyer, G., Zhong, Z. Z. (2021). Pushing Notifications as Dynamic Information Design. Marketing Science

Jerath, K., Ren, Q. (2021). Consumer Attention to Favorable and Unfavorable Product Information, and Firm Information Design. Journal of Marketing Research

Jerath, K., Ren, Q. (2022). Consumer Search and Product Returns.

Kamenica, E., Gentzkow, M. (2011). Bayesian persuasion. American Economic Review, 101(6), 2590-2615.

Ke, T. T., Lin, S., Lu, M. Y. (2022). Information design of online platforms. Available at SSRN.

Ke, T. T., Shen, Z. J. M., Villas-Boas, J. M. (2016). Search for information on multiple products. Management Science, 62(12), 3576-3603.

Ke, T. T., Villas-Boas, J. M. (2019). Optimal learning before choice. Journal of Economic Theory, 180, 383-437.

Kim, J. B., Albuquerque, P., Bronnenberg, B. J. (2010). Online demand under limited consumer search. Marketing science, 29(6), 1001-1023.

Kim, J. B., Albuquerque, P., Bronnenberg, B. J. (2017). The probit choice model under sequential search with an application to online retailing. Management Science, 63(11), 3911-3929.

Kuksov, D., Lin, Y. (2017). Signaling low margin through assortment. Management Science, 63(4), 1166-1183.

Kuksov, D., Villas-Boas, J. M. (2010). When more alternatives lead to less choice. Marketing Science, 29(3), 507-524.

Kuksov, D., Zia, M. (2021). Benefits of customer loyalty in markets with endogenous search costs. Management Science, 67(4), 2171-2190.

Liu, L., Dukes, A. (2013). Consideration set formation with multiproduct firms: The case of within-firm and across-firm evaluation costs. Management Science, 59(8), 1871-1886.

Ma, L. (2016). Only the interested learn: A model of proactive learning of product reviews. Working paper.

Mayzlin, D., Shin, J. (2011). Uninformative advertising as an invitation to search. Marketing science, 30(4), 666-685.

Moraga-González, J. L., Sándor, Z., Wildenbeest, M. R. (2021). Consumer search and prices in the automobile market.

Morozov, I. (2021). Measuring benefits from new products in markets with information frictions. Available at SSRN 3446398.

Morozov, I., Seiler, S., Dong, X., Hou, L. (2021). Estimation of preference heterogeneity in markets with costly search. Marketing Science, 40(5), 871-899.

Moscarini, G., Smith, L. (2001). The optimal level of experimentation. Econometrica, 69(6), 1629-1644.

Orlov, D., Skrzypacz, A., Zryumov, P. (2020). Persuading the principal to wait. Journal of Political Economy, 128(7), 2542-2578.

Renault, J., Solan, E., Vieille, N. (2017). Optimal dynamic information provision. Games and Economic Behavior, 104, 329-349.

Robert, J., Stahl, D. O. (1993). Informative price advertising in a sequential search model. Econometrica: Journal of the Econometric Society, 657-686.

Seiler, S. (2013). The impact of search costs on consumer behavior: A dynamic approach. Quantitative Marketing and Economics, 11(2), 155-203.

Seiler, S., Pinna, F. (2017). Estimating search benefits from path-tracking data: measurement and determinants. Marketing Science, 36(4), 565-589.

Stigler, G. J. (1961). The economics of information. Journal of political economy, 69(3), 213-225.

Ursu, R. M., Wang, Q., Chintagunta, P. K. (2020). Search duration. Marketing Science, 39(5), 849-871.

Villas-Boas, J. M. (2009). Product variety and endogenous pricing with evaluation costs. Management Science, 55(8), 1338-1346.

Villas-Boas, J. M., Yao, Y. (2021). A dynamic model of optimal retargeting. Marketing Science, 40(3), 428-458.

Wei, D. (2021). Persuasion under costly learning. Journal of Mathematical Economics, 94, 102451.

Weitzman, M. L. (1979). Optimal search for the best alternative. Econometrica: Journal of the Econometric Society, 641-654.

Wolinsky, A. (1986). True monopolistic competition as a result of imperfect information. The Quarterly Journal of Economics, 101(3), 493-511.

Yao, Y. J. (2023). Multi-attribute Search and Informative Advertising.

Yavorsky, D., Honka, E., Chen, K. (2021). Consumer search in the US auto industry: The role of dealership visits. Quantitative Marketing and Economics, 19(1), 1-52.

Zhong, W. (2022a). Optimal dynamic information acquisition. Econometrica, 90(4), 1537-1582.

Zhong, Z. (2022b). Platform search design: The roles of precision and price. Marketing Science.

Appendix

Proof of Proposition 1. We first characterize the optimal one-period strategy (providing an one-shot signal) of the sender. Analagous to section 2.3.1, the sender's problem is:

$$\max_{\lambda_0,\bar{\mu_0}} -K(\lambda_0)+p\lambda_0 \tag{P_0}$$

$$\text{s.t. } \lambda_0(\bar{\mu_0}+v_b)\geq c \tag{IR_0'}$$

$$(F_0), \lambda_0\in[0,1], \underline{\mu}_0\in[0,\mu_0)$$

We transform (P_0) into an equivalent program that is easier to analyze.

Lemma 2. *If $\mu_0 < c/v_g$, the sender does not provide information in the second period. If $\mu_0 \geq c/v_g$, (P_0) is equivalent to:*

$$\Pi_1(\mu_0) := \max -K(\lambda_0)+p\lambda_0 \tag{P_0'}$$

$$s.t.\ \lambda_0 \in \left[\frac{c}{v_g}, \frac{\mu_0-c}{-v_b}\right]$$

Proof of Lemma 2. We first show that any (λ_0,μ_0) satisfying the constraints in (P_0) also satisfy the constraints in (P_0'): $(IR_0') \Rightarrow \lambda_0 \geq \frac{c}{\bar{\mu_0}+v_b} \geq \frac{c}{v_g}$. (IR_0') & $(F_0) \Rightarrow \lambda_0 \leq \frac{\mu_0-c-\underline{\mu}_0}{-v_b-\underline{\mu}_0} \leq \frac{\mu_0-c}{-v_b}$. Thus, $\lambda_0 \in \left[\frac{c}{v_g}, \frac{\mu_0-c}{-v_b}\right]$. It is feasible for the sender to provide information in the second period iff $\left[\frac{c}{v_g}, \frac{\mu_0-c}{-v_b}\right]$ is non-empty: $\frac{c}{v_g} \leq \frac{\mu_0-c}{-v_b} \Leftrightarrow \mu_0 \geq \frac{c}{v_g}$. So, If $\mu_0 < \frac{c}{v_g}$, the sender will not provide information in the second period.

We then show that for any (λ_0,μ_0) satisfying the constraints in (P_0') and $\mu_0 \geq \frac{c}{v_g}$, we can find $\bar{\mu_0}, \underline{\mu}_0$ such that $(\lambda_0,\mu_0,\bar{\mu_0},\underline{\mu}_0)$ satisfies the constraints in (P_0). The conclusion then follows. Suppose (λ_0,μ_0) satisfies the constraints in (P_0'): $\lambda_0 \in \left[\frac{c}{v_g}, \frac{\mu_0-c}{-v_b}\right], \mu_0 \geq \frac{c}{v_g}$. Consider $\bar{\mu_0} = \frac{c}{\lambda_0} - v_b$ and $\underline{\mu}_0 = \frac{\mu_0-c+v_b\lambda_0}{1-\lambda_0}$. One can verify that $(\lambda_0,\mu_0,\bar{\mu_0},\underline{\mu}_0)$ satisfies (IR_0') & (F_0). So, we just need to show that $-v_b \leq \bar{\mu_0} \leq 1$ and $\underline{\mu}_0 \geq 0$. $\bar{\mu_0} = \frac{c}{\lambda_0} - v_b \geq -v_b$. $\lambda_0 \geq \frac{c}{v_g} \Rightarrow \bar{\mu_0} = \frac{c}{\lambda_0} - v_b \leq 1$. $\lambda_0 \leq \frac{\mu_0-c}{-v_b} \Rightarrow \mu_0 \geq c - v_b\lambda_0 \Rightarrow \underline{\mu}_0 = \frac{\mu_0-c+v_b\lambda_0}{1-\lambda_0} \geq 0$. □

Now consider the transformed program (P_0') when $\mu_0 \geq \frac{c}{v_g}$.

1. If $c \geq v_g\lambda_1^{**}$ (*i.e.* $\lambda_1^{**} \leq \frac{c}{v_g}$) and the sender provides information, then $\lambda_0^* = \frac{c}{v_g}$ due to strict concavity of the objective function. One can show that $(\lambda_0,\bar{\mu_0},\underline{\mu}_0) = (\frac{c}{v_g}, 1, \frac{\mu_0 v_g - c}{v_g - c})$ is the only feasible information structure that satisfies (IR_0') and (F_0). Thus, the sender will provide information with $(\lambda_0,\bar{\mu_0},\underline{\mu}_0) = (\frac{c}{v_g}, 1, \frac{\mu_0 v_g - c}{v_g - c})$ iff the sender surplus, $-K(\frac{c}{v_g}) + p\cdot\frac{c}{v_g}$, is positive (when it is 0, the sender is indifferent between providing information or not). Let $f(\tilde{c}) = -K(\frac{\tilde{c}}{v_g}) + p\cdot\frac{\tilde{c}}{v_g}$. We have $f(0)=0$, f is strictly concave and obtains the maximum at $\tilde{c}^* = v_g\lambda_1^{**} < c < 1$. In addition, $f(v_g) < 0$ because $\lim_{\lambda\to 1} K'(\lambda) = +\infty$. Therefore, there exists a unique $\widehat{c} \in (v_g\lambda_1^{**}, v_g)$ $s.t.$ $f(c) \begin{cases} \geq 0, & if\ 0 \leq c \leq \widehat{c} \\ < 0, & if\ c > \widehat{c} \end{cases}$. Moreover, when the sender provides

information, $\mu_0 \geq \frac{c}{v_g} \Rightarrow \widehat{c} \leq \mu_0 v_g$. So, the sender does not provide information if $c > \widehat{c}$ and provides information with $(\lambda_0^*, \bar{\mu}_0^*) = (\frac{c}{v_g}, 1)$ if $c < \widehat{c}$. The receiver surplus is 0.

2. If $c \in [\mu_0 + v_b\lambda_1^{**}, v_g\lambda_1^{**})$ $(i.e.\ \lambda_1^{**} \geq \frac{\mu_0 - c}{-v_b} > \frac{c}{v_g})$ and the sender provides information, then $\lambda_0^* = \frac{\mu_0 - c}{-v_b}$ due to strict concavity of the objective function. One can show that $(\lambda_0, \bar{\mu_0}, \underline{\mu}_0) = (\frac{\mu_0 - c}{-v_b}, \frac{-\mu_0 v_b}{\mu_0 - c}, 0)$ is the only feasible information structure that satisfies (IR_0') and (F_0). Thus, the sender will provide information with $(\lambda_0, \bar{\mu_0}, \underline{\mu}_0) = (\frac{\mu_0 - c}{-v_b}, \frac{-\mu_0 v_b}{\mu_0 - c}, 0)$ iff the sender surplus, $-K(\frac{\mu_0 - c}{-v_b}) + p \cdot \frac{\mu_0 - c}{-v_b}$, is positive. Since $-K(0) + p \cdot 0 = 0$, $\frac{\mu_0 - c}{-v_b} < \lambda_1^{**}$, and the objective function is strictly concave, the sender surplus is always strictly positive. So, the sender will always provide information. The receiver surplus is zero.

3. If $c < \mu_0 + v_b\lambda_1^{**} \wedge v_g\lambda_1^{**}$ $(i.e.\ \lambda_1^{**} \in \left(\frac{c}{v_g}, \frac{\mu_0 - c}{-v_b}\right))$, then the sender can obtain the maximum possible payoff by setting $(\lambda_0, \bar{\mu_0}) = (\lambda_1^{**}, \frac{\mu_0}{\lambda_1^{**}} \wedge 1)$.[14] Let $\underline{\mu}_0 = \begin{cases} 0, \ if\ \mu_0 \leq \lambda_1^{**} \\ \frac{\mu_0 - \lambda_1^{**}}{1 - \lambda_1^{**}}, \ if\ \mu_0 > \lambda_1^{**} \end{cases}$. One can verify that $(\lambda_0, \bar{\mu_0}, \underline{\mu}_0)$ is feasible and satisfies (IR_0') and (F_0). We have shown in the proof of Lemma 1 that the sender surplus is strictly positive. So, the sender will provide information and $(\lambda_0^*, \bar{\mu}_0^*) = (\lambda_1^{**}, \frac{\mu_0}{\lambda_1^{**}} \wedge 1)$. The receiver surplus is $\begin{cases} \mu_0 + v_b\lambda_1^{**} - c, \ if\ \mu_0 \leq \lambda_1^{**} \\ \lambda_1^{**} v_g - c, \ if\ \mu_0 > \lambda_1^{**} \end{cases} > 0$.

There are two types of iterative signals.

(a) The receiver searches regardless of the signal realization in the first period, and takes action G (B) after observing a positive (negative) signal in the second period (RHS of Figure 4).

Denote the information structure in the first period by $(\lambda_0, \bar{\mu}_0, \underline{\mu}_0)$. Denote the information structure in the second period by $(\lambda_1^p, \bar{\mu}_1^p, \underline{\mu}_1^p)$ if the receiver observes a positive signal in the first period, and by $(\lambda_1^n, \bar{\mu}_1^n, \underline{\mu}_1^n)$ if the receiver observes a negative signal in the first period. Now consder a one-period strategy $(\lambda_0', \bar{\mu}_0', \underline{\mu}_0') = (\lambda_0\lambda_1^p + (1-\lambda_0)\lambda_1^n, \frac{\lambda_0\lambda_1^p}{\lambda_0\lambda_1^p + (1-\lambda_0)\lambda_1^n}\bar{\mu}_1^p + \frac{(1-\lambda_0)\lambda_1^n}{\lambda_0\lambda_1^p + (1-\lambda_0)\lambda_1^n}\bar{\mu}_1^n, \frac{\mu_0 - \lambda_0\lambda_1^p\bar{\mu}_1^p - (1-\lambda_0)\lambda_1^n\bar{\mu}_1^n}{1 - \lambda_0\lambda_1^p - (1-\lambda_0)\lambda_1^n})$. One can check that the variables are well-defined and the beliefs are feasible. We now check the participation constraint. $\lambda_0'(\bar{\mu}_0' + v_b) = [\lambda_0\lambda_1^p + (1-\lambda_0)\lambda_1^n](\bar{\mu}_0' + v_b) = \lambda_0\lambda_1^p(\bar{\mu}_1^p + v_b) + (1-\lambda_0)\lambda_1^n(\bar{\mu}_1^n + v_b) \geq 2c \geq c$, where the first inequality comes from the first-period participation constraint for the iterative signals.

$$\begin{aligned}
\Pi_1(\mu_0) \geq& -K(\lambda_0') + p\lambda_0' \\
>& -\lambda_0 K(\lambda_1^p) - (1-\lambda_0)K(\lambda_1^n) + \lambda_0 p\lambda_1^p + (1-\lambda_0)p\lambda_1^n \text{ (convexity of } K) \\
>& -K(\lambda_0) + \lambda_0(-K(\lambda_1^p) + p\lambda_1^p) + (1-\lambda_0)(-K(\lambda_1^n) + p\lambda_1^n) \\
=& \text{ sender's payoff using the iterative signals}
\end{aligned}$$

(b) The receiver searches (takes action B) if the signal is positive (negative) in the first period, and takes action G (B) if the signal is positive (negative) in the second period (LHS of Figure 4).

[14] The notation $a \wedge b$ means the minimum of a and b.

If $c \geq v_g\lambda_1^{**}$, or $c < v_g\lambda_1^{**}$ and $\mu_1 \leq c - v_b\lambda_1^{**}$, the expected receiver surplus in the second period is 0. The receiver incurs search cost without any immediate benefit in the first period under iterative signals. The expected receiver surplus in the first period is strictly negative if he searches. Therefore, he will not search, and iterative signals are not feasible. Now we consider the case in which $c < v_g\lambda_1^{**}$ and $\mu_1 > c - v_b\lambda_1^{**}$. The sender's problem when she uses such iterative signals is:

$$\Pi_{iter}(\mu_0) := \max -K(\lambda_0) + \lambda_0 \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \tag{P_{iter}}$$

$$\text{s.t. } \lambda_1(\bar{\mu_1} + v_b) \geq \frac{1+\lambda_0}{\lambda_0}c \tag{$IR_{0,iter}$}$$

$$\lambda_1(\bar{\mu_1} + v_b) \geq c \tag{$IR_{1,iter}$}$$

$$(F_0), (F_1), \mu_1 = \bar{\mu_0}, \lambda_1 = \lambda_1^{**}$$

Note that $(IR_{0,iter})$ implies (IR_1^{iter}) and $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}})$ satisfies (F_1). $\forall \mu_1 \geq \lambda_1^{**}$, the optimal second-period strategy is always $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, 1)$. Therefore, choosing μ_1 above λ_1^{**} does not increase the second-period sender's payoff or relax the first-period constraints. So, we can restrict μ_1 to be less than or equal to λ_1^{**}.

i) $\mu_0 \geq c - v_b\lambda_1^{**}$: $\Pi_1(\mu_0) = -K(\lambda_1^{**}) + p\lambda_1^{**} > \Pi_{iter}(\mu_0)$.

ii) $\mu_0 < c - v_b\lambda_1^{**}$: $\Pi_1(\mu_0) = -K(\frac{\mu_0 - c}{-v_b}) + \frac{(\mu_0 - c)p}{-v_b}$

$$(F_1) \text{ \& } (IR_{0,iter}) \Rightarrow \lambda_0 \geq \frac{c}{\mu_1 - c + v_b\lambda_1^{**}} (\Rightarrow \mu_1 \geq \frac{1+\lambda_0}{\lambda_0}c - v_b\lambda_1^{**}) \tag{1}$$

$$\overset{\mu_1 \leq \lambda_1^{**}}{\geq} \frac{c}{v_g\lambda_1^{**} - c}$$

$$(F_0) \Rightarrow \lambda_0 = \frac{\mu_0 - \underline{\mu}_0}{\mu_1 - \underline{\mu}_0} \leq \frac{\mu_0}{\mu_1} \overset{(1)}{\leq} \frac{\mu_0}{\frac{1+\lambda_0}{\lambda_0}c - v_b\lambda_1^{**}} \Rightarrow \lambda_0 \leq \frac{\mu_0 - c}{c - v_b\lambda_1^{**}}$$

A necessary condition for λ_0 to be well-defined is:

$$\frac{c}{v_g\lambda_1^{**} - c} \leq \frac{\mu_0 - c}{c - v_b\lambda_1^{**}} \Leftrightarrow \mu_0 \geq \frac{\lambda_1^{**}c}{v_g\lambda_1^{**} - c} (> \frac{c}{v_g})$$

Therefore, it is feasible for the sender to provide a one-period signal whenever it is feasible to provide iterative signals. Define $\bar{\Pi}_{iter}(\mu_0) := \max_{0 \leq \lambda_0 \leq \frac{\mu_0 - c}{c - v_b\lambda_1^{**}}} -K(\lambda_0) + \lambda_0 \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$. One can see that $\bar{\Pi}_{iter}(\mu_0) \geq \Pi_{iter}(\mu_0)$. Let $\lambda_0^*(\mu_0) := \arg\max_{0 \leq \lambda_0 \leq \frac{\mu_0 - c}{c - v_b\lambda_1^{**}}} -K(\lambda_0) + \lambda_0 \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$.

We want to show:

$$\bar{\Pi}_{iter}(\mu_0) < \Pi_1(\mu_0) \Leftrightarrow -K(\lambda_0^*(\mu_0)) + \lambda_0^*(\mu_0) \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] < -K(\frac{\mu_0 - c}{-v_b}) + \frac{\mu_0 - c}{-v_b}p \tag{2}$$

$$\text{Notice that } \frac{d}{d\lambda_0}\left\{-K(\lambda_0) + \lambda_0 \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]\right\} = -K'(\lambda_0) - K(\lambda_1^{**}) + p\lambda_1^{**}$$

$$\Rightarrow \lambda_0^*(\mu_0) = \begin{cases} \frac{\mu_0 - c}{c - v_b \lambda_1^{**}} & if\ \mu_0 \le \mu_0^t \\ \frac{\mu_0^t - c}{c - v_b \lambda_1^{**}} & if\ \mu_0 > \mu_0^t \end{cases}$$

, where $\lambda_0^t > 0$ is defined by $-K'(\lambda_0^t) - K(\lambda_1^{**}) + p\lambda_1^{**} = 0$, $\mu_0^t = \lambda_0^t(c - v_b\lambda_1^{**}) + c$.

When $\mu_0 \le \mu_0^t$, $\begin{cases} \lambda_0^*(\mu_0) = \frac{\mu_0 - c}{c - v_b\lambda_1^{**}} > \frac{\mu_0 - c}{-v_b} \Rightarrow -K(\lambda_0^*(\mu_0)) < -K(\frac{\mu_0 - c}{-v_b}) \\ \lambda_0^*(\mu_0) p\lambda_1^{**} = \frac{\mu_0 - c}{c - v_b\lambda_1^{**}} p\lambda_1^{**} < \frac{\mu_0 - c}{-v_b} p \\ -\lambda_0^*(\mu_0) K(\lambda_1^{**}) < 0 \end{cases}$ $\Rightarrow$ (2) holds, where

the first inequality holds because $-v_b > \mu_1 > c - v_b\lambda_1^{**}$.

When $\mu_0 > \mu_0^t$, $\bar{\Pi}_{iter}(\mu_0) = \bar{\Pi}_{iter}(\mu_0^t) < \Pi_1(\mu_0^t) < \Pi_1(\mu_0)$. So, (2) holds.

Thus, $\Pi_1(\mu_0) > \bar{\Pi}_{iter}(\mu_0)$ for any μ_0 such that iterative signals are feasible. □

Proof of Lemma 1. λ_0^{**} and λ_1^{**} are determined by the first order conditions: $-K'(\lambda_1^{**}) + p = 0$ and $-K'(\lambda_0^{**}) + p + K(\lambda_1^{**}) - p\lambda_1^{**} = 0$. $-K(0) + p \cdot 0 = 0\ \&\ -K'(\lambda) + p > 0$ for small $\lambda \Rightarrow -K(\lambda_1^{**}) + p\lambda_1^{**} > 0$. Therefore, $-K'(\lambda_0^{**}) + p + K(\lambda_1^{**}) - p\lambda_1^{**} = 0$ implies that $-K'(\lambda_0^{**}) + p > 0 = -K'(\lambda_1^{**}) + p, \Rightarrow K'(\lambda_0^{**}) < K'(\lambda_1^{**}) \Rightarrow \lambda_0^{**} < \lambda_1^{**}$. $-K(0) + p \cdot 0 + (1-0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] > 0$ and strict concavity (w.r.t. λ_0) of the objective function imply that $-K(\lambda_0^{**}) + p\lambda_0^{**} + (1 - \lambda_0^{**})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] > 0$. □

Proof of Proposition 4. It can be proved in the same way as in the proof of Proposition 1 where we derive the optimal one-period strategy of the sender. □

Proof of Proposition 5. **Road map for the proof:**

(1) $v_g\lambda_1^{**} \le c < \widehat{c}$

Lemma 3

↓

Lemma 4

↓

optimal two-period strategy

↓

optimal strategy

(2) $c < v_g\lambda_1^{**}$

(P_{2S_+}): Proposition 9 ← Lemma 5

(P_{2S_0}): Proposition 10 ← Lemma 7 ← Lemma 6

↓

Comparing one-period and two-period sender's payoff ← Lemma 8 & Proposition 1

Detailed proof:

(1) $v_g\lambda_1^{**} \le c < \widehat{c}$

If the sender provides information in both periods, the sender's constrained program is:

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right] \tag{P_{2H}}$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \geq \frac{c}{v_g}$$

We first transform (P_{2H}) into an equivalent program that is easier to analyze.

Lemma 3. *Suppose $v_g\lambda_1^{**} \leq c < \widehat{c}$. If $\mu_{0,1} \leq \mu_0 < \frac{2v_g-c}{(v_g)^2}c$, the sender provides information in one period. If $\mu_0 \geq \frac{2v_g-c}{(v_g)^2}c$, ($P_{2H}$) is equivalent to:*

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right] \tag{P'_{2H}}$$
$$s.t. \;\; \lambda_0 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - (1+v_g)c}{-v_b v_g - c}\right]$$

Proof of Lemma 3. The proof of the equivalence between (P_{2H}) and (P'_{2H}) is similar to that of Lemma 2. It is feasible for the sender to provide information at both periods if and only if the domain of λ_1 is non-empty: $\frac{c}{v_g} \leq \frac{v_g\mu_0-(2-p)c}{pv_g-c} \Leftrightarrow \mu_0 \geq \frac{2v_g-c}{(v_g)^2}c$. □

Denote the optimal λ_0 without constraints by $\lambda_{0,H}^{**}$. $\lambda_{0,H}^{**} = \arg\max_{\lambda_0} -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right]$. The following lemma summarizes the relative size of $\lambda_{0,H}^{**}$, λ_0^{**}, and $\frac{c}{v_g}$.

Lemma 4. $0 < \lambda_{0,H}^{**} < \lambda_1^{**} \leq \frac{c}{v_g}$.

Proof of Lemma 4. $\lambda_1^{**} < \frac{c}{v_g}$ is the assumption. F.O.C $\Rightarrow K'(\lambda_{0,H}^{**}) = p + K(\frac{c}{v_g}) - \frac{cp}{v_g}$. From Lemma 1, $K'(\lambda_1^{**}) = p$. $-K(\frac{c}{v_g}) + \frac{cp}{v_g} > 0$ when $c < \widehat{c}$. So, $K'(\lambda_{0,H}^{**}) < K'(\lambda_1^{**}) \Rightarrow \lambda_{0,H}^{**} < \lambda_1^{**}$. $-K'(\lambda_{0,H}^{**}) + p + K(\frac{c}{v_g}) - \frac{cp}{v_g} = 0 \Rightarrow K'(\lambda_{0,H}^{**}) = p + K(\frac{c}{v_g}) - \frac{cp}{v_g} > p - \frac{cp}{v_g} = (1-\frac{c}{v_g})p > 0$, where the last inequality follows from the assumption that $c < v_g$. Thus, $\lambda_{0,H}^{**} > 0$. □

When it is feasible for the sender to provide information in both periods, $\mu_0 \geq \frac{2v_g-c}{(v_g)^2}c$, Lemma 4 and strict concavity of the objective function imply that the optimal two-period strategy of the sender is $(\lambda_t^*, \bar{\mu}_t{}^*) = (\frac{c}{v_g}, 1), t = 0, 1$. The sender surplus is $(2-\frac{c}{v_g})\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right] > -K(\frac{c}{v_g}) + \frac{cp}{v_g}$, the sender surplus of the optimal one-period strategy. Therefore, the sender will always provide information in both periods as long as it is feasible.

(2) $c < v_g\lambda_1^{**}$

If the sender provides information in both periods, we first show that we can restrict the domain of μ_1 to be $\leq \lambda_1^{**}$. The intuition is that the optimal second-period strategy is always $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, 1)$, $\forall \mu_1 \geq \lambda_1^{**}$. Therefore, choosing μ_1 above λ_1^{**} does not increase the second-period sender's payoff or relax the first-period constraints. Formally, when $\lambda_1^{**} \leq \mu_1 < \mu_0$, the sender's constrained program is:

$$\begin{aligned}&\max -K(\lambda_0)+p\lambda_0+(1-\lambda_0)\left[-K(\lambda_1^{**})+p\lambda_1^{**}\right]\\ \text{s.t. }\ &\lambda_0(\bar{\mu_0}+v_b)+(1-\lambda_0)[\lambda_1^{**}v_g-c]\geq c \qquad (\tilde{IR}_0')\\ &(F_0),\mu_1\in[\lambda_1^{**},\mu_0)\end{aligned}$$

$(\tilde{IR}_0')$ & $(F_0)\Rightarrow \lambda_0\leq \frac{-2c+\mu_0-\mu_1+v_g\lambda_1^{**}}{v_g\lambda_1^{**}-c-v_b-\mu_1}\leq \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$ ("=" when $\mu_1=\lambda_1^{**}$). $(F_0)\Rightarrow\lambda_0\geq\frac{\mu_0-\mu_1}{1-\mu_1}$. The domain of λ_0 is non-empty iff $\frac{\mu_0-\mu_1}{1-\mu_1}\leq\frac{-2c+\mu_0-\mu_1+v_g\lambda_1^{**}}{v_g\lambda_1^{**}-c-v_b-\mu_1}\Leftrightarrow \mu_1\leq\frac{-2c+v_g\lambda_1^{**}+\mu_0[1-v_g\lambda_1^{**}+c+v_b]}{v_g-c}$. Therefore, smaller μ_1 means it is more likely for the domain of λ_0 to be non-empty and larger upper bound of λ_0. So, the optimal μ_1 will never $\in(\lambda_1^{**},\mu_0)$. Hence, the optimal strategy in the second period is $(\lambda_1^*,\bar{\mu}_1^*)=\begin{cases}(\lambda_1^{**},\frac{\mu_1}{\lambda_1^{**}}) & ,\ if\ \mu_1\in(c-v_b\lambda_1^{**},\lambda_1^{**}]\\ (\frac{\mu_1-c}{-v_b},\frac{-\mu_1v_b}{\mu_1-c}) & ,\ if\ \mu_1\in[\frac{c}{v_g},c-v_b\lambda_1^{**}]\end{cases}$ and the constrained program of the entire game is either[15]:

$$\begin{aligned}&\max -K(\lambda_0)+p\lambda_0+(1-\lambda_0)\left[-K(\lambda_1^{**})+p\lambda_1^{**}\right] \qquad (P_{2S_+})\\ \text{s.t. }\ &(IR_0),(F_0),\mu_1\in[c-v_b\lambda_1^{**},\lambda_1^{**}]\end{aligned}$$

or:

$$\begin{aligned}&\max -K(\lambda_0)+p\lambda_0+(1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b})+\frac{(\mu_1-c)p}{-v_b}\right] \qquad (P_{2S_0})\\ \text{s.t. }\ &(IR_0),(F_0),\mu_1\in[\frac{c}{v_g},c-v_b\lambda_1^{**}]\end{aligned}$$

We consider the two programs above separately, and then compare the corresponding local solutions to pin down the global solution.

1. S_+ strategy (solution to (P_{2S_+}))

Proposition 9. *Suppose $c<v_g\lambda_1^{**}$ and $\mu_0<\widehat{\mu_0}=2c-v_b\lambda_1^{**}-[c+(1-\lambda_1^{**})v_b]\lambda_0^{**}$. If $\mu_0>2c-v_b\lambda_1^{**}$ and $\mu_0\geq\frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$, (P_{2S_+}) is feasible with the following solution. $\lambda_0^*=\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$; $\bar{\mu}_0^*=\begin{cases}\frac{(v_b-c)(c-v_b\lambda_1^{**})-v_b\mu_0}{\mu_0-2c+v_b\lambda_1^{**}}\in(-v_b,1) & ,if\ \widehat{\mu_1}(\mu_0)<c-v_b\lambda_1^{**}\\ 1 & ,if\ \widehat{\mu_1}(\mu_0)\geq c-v_b\lambda_1^{**}\end{cases}$; $(\lambda_1^*,\bar{\mu}_1^*)=(\lambda_1^{**},\frac{\mu_1}{\lambda_1^{**}})$; $\mu_1^*=\widehat{\mu_1}(\mu_0)\vee c-v_b\lambda_1^{**}$, where $\widehat{\mu_1}(\mu_0)=\frac{2c-v_b\lambda_1^{**}-(1+c-v_b\lambda_1^{**}+v_b)\mu_0}{c-v_b-\mu_0}$. The receiver gets negative surplus in the first period, positive surplus in the second period, and zero total surplus.*

Proof of Proposition 9. We first transform (P_{2S_+}) into an equivalent program that is easier to analyze.

Lemma 5. *Suppose $c<v_g\lambda_1^{**}$ and $\mu_0\geq\mu_{0,1}$. If $\mu_0\leq 2c-v_b\lambda_1^{**}$ or $\mu_0<\frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$, the sender provides information in one period. If $\mu_0>2c-v_b\lambda_1^{**}$ and $\mu_0\geq\frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$, (P_{2S_+}) is equivalent to:*

[15] We include $\mu_1=c-v_b\lambda_1^{**}$ in (P_{2S_+}) as well to simplify the exposition.

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \qquad (P''_{2S_+})$$
$$s.t.\ \lambda_0 \in \left(0, \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**}) - c}\right]$$

Proof of Lemma 5. We first show that, if $\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$, ($P_{2S_+}$) is equivalent to:

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \qquad (P'_{2S_+})$$
$$\text{s.t. } \lambda_0 \in \left[\frac{\mu_0 - \mu_1}{1-\mu_1}, \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**}) - c} \wedge \frac{\mu_0 - \mu_1}{-v_b - \mu_1}\right]$$
$$\mu_1 \in [c - v_b\lambda_1^{**} \vee \widehat{\mu_1}(\mu_0), \lambda_1^{**}]$$
$$\text{, where } \widehat{\mu_1}(\mu_0) = \frac{2c - v_b\lambda_1^{**} - (1 + c - v_b\lambda_1^{**} + v_b)\mu_0}{c - v_b - \mu_0}$$

$(F_0) \Rightarrow \lambda_0 = \frac{\mu_0 - \mu_1}{\bar{\mu_0} - \mu_1} \in \left[\frac{\mu_0-\mu_1}{1-\mu_1}, \frac{\mu_0-\mu_1}{-v_b-\mu_1}\right]$. (IR_0) & $(F_0) \Rightarrow \lambda_0 \leq \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$. For λ_0 to be positive, we need $\mu_0 > 2c - v_b\lambda_1^{**}$. The domain of λ_0 is non-empty iff $\frac{\mu_0-\mu_1}{1-\mu_1} \leq \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c} \Leftrightarrow \mu_1 \geq \widehat{\mu_1}(\mu_0)$. For $\mu_1 \leq \lambda_1^{**}$, we need $\widehat{\mu_1}(\mu_0) \leq \lambda_1^{**} \Leftrightarrow \mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$. We also have that $\mu_1 = \underline{\mu}_0 < \mu_0$. Thus, the constraints in (P_{2S_+}) imply the constraints in (P'_{2S_+}).

For any (λ_0, μ_1) satisfying the constraints in (P'_{2S_+}), consider $(\lambda_0, \mu_1, \bar{\mu_0} = \frac{\mu_0 - (1-\lambda_0)\mu_1}{\lambda_0}, \bar{\mu_1} = \frac{\mu_1}{\lambda_1^{**}} \wedge 1, \underline{\mu}_1 = \frac{\mu_1 - \lambda_1^{**}\bar{\mu_1}}{1-\lambda_1^{**}})$. (IR_0) & (F_0) are satisfied by construction. $\bar{\mu_0} = \frac{\mu_0-(1-\lambda_0)\mu_1}{\lambda_0} > \frac{\mu_0-(1-\lambda_0)\mu_0}{\lambda_0} = \mu_0$. $\bar{\mu_0} = \frac{\mu_0-(1-\lambda_0)\mu_1}{\lambda_0} \leq \frac{\mu_0 - (1-\frac{\mu_0-\mu_1}{1-\mu_1})\mu_1}{\frac{\mu_0-\mu_1}{1-\mu_1}} = 1$. One can verify that $\bar{\mu_1} \in (-v_b, 1]$, $\underline{\mu}_1 \in [0, -v_b)$. Therefore, the $(\lambda_0, \mu_1, \bar{\mu_0}, \bar{\mu_1}, \underline{\mu}_1)$ we constructed satisfies all the constraints in (P_{2S_+}) and is feasible. Therefore, the two programs are equivalent.

We then show that (P'_{2S_+}) is equivalent to (P''_{2S_+}). It is clear that the constraints in (P'_{2S_+}) imply the constraints in (P''_{2S_+}). We now show that for any $\lambda_0 \in \left(0, \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}\right]$, we can find a feasible (λ_0, μ_1) that satisfies the constraints in (P'_{2S_+}). Since $\lambda_0 = \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$ maximizes the objective function among $\lambda_0 \in \left(0, \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}\right]$ when $\mu_0 < \widehat{\mu_0}$, we only need to verify (by construction) that $\lambda_0 = \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$ can be obtained.

i) $\widehat{\mu_1}(\mu_0) < c - v_b\lambda_1^{**}$: Consider $\mu_1 = c - v_b\lambda_1^{**}, \lambda_0 = \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, \bar{\mu_0} = \frac{\mu_0-(1-\lambda_0)\mu_1}{\lambda_0} = \frac{(v_b-c)(c-v_b\lambda_1^{**})-v_b\mu_0}{\mu_0-2c+v_b\lambda_1^{**}}$. By construction, (IR_0) and (F_0) are satisfied; μ_1's constraints are also satisfied. So, we just need to verify that $\bar{\mu_0} \in (p, 1)$. $\bar{\mu_0} < 1 \Leftrightarrow (v_b - c)(c - v_b\lambda_1^{**}) - v_b\mu_0 < \mu_0 - 2c + v_b\lambda_1^{**} \Leftrightarrow \widehat{\mu_1}(\mu_0) < c - v_b\lambda_1^{**}$, which is the assumption. $\bar{\mu_0} > -v_b \Leftrightarrow c < -v_b(1-\lambda_1^{**})$, which holds because $\mu_0 > 2c - v_b\lambda_1^{**} \Rightarrow c < \frac{1}{2}(\mu_0 + v_b\lambda_1^{**}) \leq \mu_0 + v_b\lambda_1^{**} < -v_b + v_b\lambda_1^{**} = -v_b(1-\lambda_1^{**})$.

ii) $\widehat{\mu_1}(\mu_0) \geq c - v_b\lambda_1^{**}$: Consider $\lambda_0 = \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, \bar{\mu_0} = 1, \mu_1 = \frac{\mu_0-\lambda_0\bar{\mu_0}}{1-\lambda_0} = \widehat{\mu_1}(\mu_0)$. By construction, (IR_0) and (F_0) are satisfied; μ_1's constraints are also satisfied. So, we just need to verify that $\mu_1 = \widehat{\mu_1}(\mu_0) \in [c - v_b\lambda_1^{**}, \lambda_1^{**}]$. $\widehat{\mu_1}(\mu_0) \geq c - v_b\lambda_1^{**}$ is the assumption.

$\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c} \Rightarrow \widehat{\mu_1}(\mu_0) \leq \lambda_1^{**}$. $\square$

When $\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c} \geq \lambda_0^{**} (\Leftrightarrow \mu_0 \geq \widehat{\mu_0})$, the optimal λ_0 is λ_0^{**}. When $\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c} < \lambda_0^{**}$, the optimal λ_0 is $\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$ due to strict concavity of the objective function (denote it by $J(\lambda_0)$). Since the one-period optimal sender surplus is $-K(\lambda_1^{**}) + p\lambda_1^{**} = J(0)$, $J(\cdot)$ is strictly concave and obtains the unique maximum value at $\lambda_0^{**} > \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$, we have $J(\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}) > J(0)$. So, the sender always provides information in both periods when it is feasible ($\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$). We will use this observation in the later proofs. According to the proof of Lemma 5, the receiver always gets zero surpluses when $\mu_0 < \widehat{\mu_0}$. $\mu_1 = \widehat{\mu_1}(\mu_0) \vee c - v_b\lambda_1^{**}$ is the smallest μ_1 that supports $\lambda_0^* = \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$, which gives the receiver the largest surplus in the first period. So, $\mu_1^* = \widehat{\mu_1}(\mu_0) \vee c - v_b\lambda_1^{**}$.

$$(F_0) \Rightarrow \bar{\mu}_0^* = \frac{\mu_0-(1-\lambda_0)\mu_1}{\lambda_0} = \begin{cases} \frac{(v_b-c)(c-v_b\lambda_1^{**})-v_b\mu_0}{\mu_0-2c+v_b\lambda_1^{**}} \in (p,1) & ,if\ \widehat{\mu_1}(\mu_0) < c - v_b\lambda_1^{**} \\ 1 & ,if\ \widehat{\mu_1}(\mu_0) \geq c - v_b\lambda_1^{**} \end{cases}$$ $\square$

2. S_0 strategy (solution to (P_{2S_0}))

Proposition 10. *Suppose $c \leq v_g\lambda_1^{**}$. When $\mu_0 \geq \frac{2v_g-c}{(v_g)^2}c$, ($P'_{2S_0}$) is feasible. λ_0^*, λ_1^*, and μ_1^* are continuous and increase in μ_0, while $\bar{\mu}_0^*$ and $\bar{\mu}_1^*$ are continuous and decrease in μ_0, in the solution to (P'_{2S_0}). The receiver gets zero surplus at each period.*

Proof of Proposition 10. We first transform (P_{2S_0}) into an equivalent program that is easier to analyze.

Lemma 6. *Suppose $c < v_g\lambda_1^{**}$. If $\mu_{0,1} \leq \mu_0 < \frac{2v_g-c}{(v_g)^2}c$, the sender provides information in one period. If $\mu_0 \geq \frac{2v_g-c}{(v_g)^2}c$, ($P_{2S_0}$) is equivalent to:*

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right] \tag{P'_{2S_0}}$$
$$s.t.\ \lambda_0 \in \left[\frac{\mu_0-\mu_1}{1-\mu_1}, \frac{\mu_0-\mu_1-c}{-v_b-\mu_1}\right]$$
$$\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c}\right]$$

Proof of Lemma 6. Using the same argument as the proof of Lemma 5, one can show that (P_{2S_0}) is equivalent to:

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right] \tag{P''_{2S_0}}$$
$$\text{s.t. } \lambda_0 \in \left[\frac{\mu_0-\mu_1}{1-\mu_1}, \frac{\mu_0-\mu_1-c}{-v_b-\mu_1}\right]$$
$$\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c} \wedge c - v_b\lambda_1^{**}\right]$$

We just need to show that (P''_{2S_0}) is equivalent to (P'_{2S_0}). If $\frac{v_g\mu_0-c}{v_g-c} \leq c-v_b\lambda_1^{**}$, μ_1's constraint becomes $\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c}\right]$. So, ($P''_{2S_0}$) is equivalent to ($P'_{2S_0}$). If $\frac{v_g\mu_0-c}{v_g-c} > c - v_b\lambda_1^{**}$, denote the solution to (P'_{2S_0}) by (λ_0, μ_1).

(a) $(\lambda_0^{**}, \lambda_1^{**})$ can be obtained ($\lambda_1 = \lambda_1^{**} \Leftrightarrow \mu_1 = c - v_b\lambda_1^{**}$) in ($P''_{2S_0}$)
$\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c} \wedge c - v_b\lambda_1^{**}\right]$ is equivalent to $\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c}\right]$, as the optimal μ_1 under the latter (relaxed) constraint will be $c - v_b\lambda_1^{**}$. So, (P''_{2S_0}) is equivalent to (P'_{2S_0}).

(b) $(\lambda_0^{**}, \lambda_1^{**})$ can not be obtained in (P''_{2S_0})
Suppose $\mu_1 > c - v_b\lambda_1^{**}$. If $\lambda_0 > \frac{\mu_0-\mu_1}{1-\mu_1}$, consider $(\lambda_0' = \lambda_0, \mu_1' = \mu_1 - \varepsilon)$. For small enough ε, it is feasible and gives the sender a strictly higher payoff. A contradiction! If $\lambda_0 = \frac{\mu_0-\mu_1}{1-\mu_1}$ instead, we have $\frac{\mu_0-\mu_1}{1-\mu_1} \geq \lambda_0^{**}$. A contradiction!
Therefore, $\mu_1 \leq c - v_b\lambda_1^{**}$ and thus (P''_{2S_0}) is equivalent to (P'_{2S_0}).

In sum, (P''_{2S_0}) is equivalent to (P'_{2S_0}). □

Lemma 7. *Suppose $c < v_g\lambda_1^{**}$. For $\mu_0 < \widehat{\mu_0}, \lambda_0$ is binding at the upper bound in the solution to (P'_{2S_0}).*

Proof of Lemma 7. To solve (P'_{2S_0}), we consider several cases.

i) $\lambda_0 \leq \frac{\mu_0-\mu_1-c}{-v_b-\mu_1}$ is binding and μ_1's constraints are not binding.
The Lagrangian is $\mathcal{L} = -K(\lambda_0)+p\lambda_0+(1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right]+\eta\left(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1} - \lambda_0\right)$
$s.t.\ \eta \geq 0, \eta\left(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1} - \lambda_0\right) = 0.$

$$\text{F.O.C.} \Rightarrow \begin{cases} -K'(\lambda_0) + p + K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{v_b} - \eta = 0 \\ (1-\lambda_0)\left[K'(\frac{\mu_1-c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right] + \eta \cdot \frac{\mu_0+v_b-c}{(v_b+\mu_1)^2} = 0 \end{cases}$$

Plug in $\lambda_0 = \frac{\mu_0-\mu_1-c}{-v_b-\mu_1}$. Dividing the second equality by $\frac{\mu_0+v_b-c}{(v_b+\mu_1)^2}$ and comparing with the first equality, we obtain:

$$\begin{aligned} \eta &= -(v_b+\mu_1)\left[K'(\frac{\mu_1-c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right] \\ &= -K'(\lambda_0) + p + K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{v_b} \\ \Rightarrow\ & K(\frac{\mu_1-c}{-v_b}) + \frac{v_b+\mu_1}{v_b}K'(\frac{\mu_1-c}{-v_b}) - K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) - \frac{cp}{v_b} = 0 \qquad (*) \end{aligned}$$

$\frac{\partial}{\partial\mu_1}\left[K(\frac{\mu_1-c}{-v_b}) + \frac{v_b+\mu_1}{v_b}K'(\frac{\mu_1-c}{-v_b})\right] = -\frac{v_b+\mu_1}{(v_b)^2}K''(\frac{\mu_1-c}{-v_b}) > 0$. So, the sum of the first two terms of the LHS of (*) strictly increases in μ_1. $\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}$ strictly decreases in μ_1, $K'(\cdot)$ strictly increases in μ_1. So, $-K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})$ strictly increases in μ_1. Thus, the LHS of (*) strictly increases in μ_1. When μ_0 increases, the LHS of (*) is strictly negative if μ_1 is unchanged. Therefore, μ_1 also has to increase. So, the sum of the first two terms of the LHS of (*) increases. As a result, the third term, $-K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) = -K'(\lambda_0)$ has to decrease strictly. So, λ_0 has to increase strictly. In sum, the optimal λ_0 and μ_1 are strictly increasing in μ_0.

ii) $\mu_1 \leq \frac{v_g\mu_0-c}{v_g-c}$ is binding.

When $\mu_1 = \frac{v_g\mu_0-c}{v_g-c}$, $\lambda_0 \in \{\frac{c}{v_g}\}$. So, λ_0 is binding at the upper bound.

iii) $\mu_1 \geq \frac{c}{v_g}$ is binding and λ_0 is not binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right] + \eta\left(\mu_1 - \frac{c}{v_g}\right)$ *s.t.* $\eta \geq 0, \eta\left(\mu_1 - \frac{c}{v_g}\right) = 0$.

$$\text{F.O.C.} \Rightarrow \begin{cases} -K'(\lambda_0) + p + K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{v_b} = 0 \\ (1-\lambda_0)\left[K'(\frac{\mu_1-c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right] + \eta = 0 \end{cases}$$

The second equality $\Rightarrow \eta = -\frac{1-\lambda_0}{v_b}\left[K'(\frac{\mu_1-c}{-v_b}) - p\right] \overset{c<v_g\lambda_1^{**}}{<} -\frac{1-\lambda_0}{v_b}\left[K'(\lambda_1^{**}) - p\right] = 0$. But $\eta \geq 0$. A contradiction! So, this case cannot happen.

iv) $\mu_1 \geq \frac{c}{v_g}$ is binding and $\lambda_0 \geq \frac{\mu_0-\mu_1}{1-\mu_1}$ is binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right] + \eta\left(\mu_1 - \frac{c}{v_g}\right) + \xi\left(\lambda_0 - \frac{\mu_0-\mu_1}{1-\mu_1}\right)$ *s.t.* $\eta \geq 0, \eta\left(\mu_1 - \frac{c}{v_g}\right) = 0, \xi \geq 0, \xi\left(\lambda_0 - \frac{\mu_0-\mu_1}{1-\mu_1}\right) = 0$.

$$\text{F.O.C.} \Rightarrow \begin{cases} -K'(\lambda_0) + p + K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{v_b} + \xi = 0 \\ (1-\lambda_0)\left[K'(\frac{\mu_1-c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right] + \eta + \xi\frac{1-\mu_0}{(1-\mu_1)^2} = 0 \end{cases}$$

Similar to the previous case, the LHS of the second equality > 0. A contradiction!

v) $\lambda_0 \geq \frac{\mu_0-\mu_1}{1-\mu_1}$ is binding and μ_1 is not binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right] + \xi\left(\lambda_0 - \frac{\mu_0-\mu_1}{1-\mu_1}\right)$ *s.t.* $\xi \geq 0, \xi\left(\lambda_0 - \frac{\mu_0-\mu_1}{1-\mu_1}\right) = 0$.

$$\text{F.O.C.} \Rightarrow \begin{cases} -K'(\lambda_0) + p + K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b} + \xi = 0 \\ (1-\lambda_0)\left[K'(\frac{\mu_1-c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right] + \xi\frac{1-\mu_0}{(1-\mu_1)^2} = 0 \end{cases}$$

Similar to the previous case, the LHS of the second equality > 0. A contradiction!

vi) both λ_0 and μ_1 are not binding.

The solution is the unconstrained optimal solution $(\lambda_0^{**}, \lambda_1^{**})$. But we have assumed that it is not feasible.

i) to vi) finish the proof of Lemma 7. ☐

According to Lemma 7, if $\mu_0 \geq \frac{2v_g-c}{(v_g)^2}c$, ($P'_{2S_0}$) is equivalent to:

$$\max -K(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) + p\,\frac{\mu_0-\mu_1-c}{-v_b-\mu_1} + (1-\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right] \qquad (P'''_{2S_0})$$
$$\text{s.t. } \mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c}\right]$$

The first order derivative of the objective function w.r.t. μ_1 is:

$$
\begin{aligned}
D(\mu_0,\mu_1):=&\frac{\partial}{\partial\mu_1}\left\{-K(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})+p\ \frac{\mu_0-\mu_1-c}{-v_b-\mu_1}+(1-\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})\left[-K(\frac{\mu_1-c}{-v_b})+\frac{(\mu_1-c)p}{-v_b}\right]\right\}\\
=&\frac{\mu_0+v_b-c}{(\mu_1+v_b)^2}\left[K(\frac{\mu_1-c}{-v_b})+\frac{v_b+\mu_1}{v_b}K'(\frac{\mu_1-c}{-v_b})-K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})-\frac{cp}{v_b}\right]
\end{aligned}
$$

The first term of $D(\mu_0,\mu_1)$, $\frac{\mu_0+v_b-c}{(\mu_1+v_b)^2}$, is always strictly negative. The second term, $K(\frac{\mu_1-c}{-v_b})+\frac{v_b+\mu_1}{v_b}K'(\frac{\mu_1-c}{-v_b})-K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})-\frac{cp}{v_b}$, is the LHS of (*), which has been shown to be strictly increasing in μ_1 in the proof of Lemma 7. One can see that $D(\mu_0,\mu_1)$ is strictly negative when μ_1 is large. Thus, $D(\mu_0,\mu_1)$ is always negative or positive for μ_1 small and negative for μ_1 large. Let $\mu_1^{**}(\mu_0)$ be the cutoff value such that $D(\mu_0,\mu_1)\geq 0$ for $\mu_1\leq\mu_1^{**}(\mu_0)$ and $D(\mu_0,\mu_1)\leq 0$ for $\mu_1\geq\mu_1^{**}(\mu_0)$ ($\mu_1^{**}(\mu_0):=-\infty$ if $D(\mu_0,\mu_1)$ is always negative). Since $\mu_1\in\left[\frac{c}{v_g},\frac{v_g\mu_0-c}{v_g-c}\right]$, the optimal $\mu_1^*(\mu_0)=\left[\frac{c}{v_g}\vee\mu_1^{**}(\mu_0)\right]\wedge\frac{v_g\mu_0-c}{v_g-c}$. One can see that we can define $\tilde{\mu}_1^{**}(\mu_0):=\begin{cases}\frac{c}{v_g},\ if\ D(\mu_0,\mu_1)\text{ is always negative}\\ \mu_1^{**}(\mu_0),\ \text{otherwise}\end{cases}$. $\tilde{\mu}_1^{**}(\mu_0)\in(-\infty,+\infty)$ and $\mu_1^*(\mu_0)=\left[\frac{c}{v_g}\vee\tilde{\mu}_1^{**}(\mu_0)\right]\wedge\frac{v_g\mu_0-c}{v_g-c}$. Since $\tilde{\mu}_1^{**}(\mu_0)$ is continuous in μ_0, $\mu_1^*(\mu_0)$ is also continuous in μ_0. It then implies that $\lambda_0^*(\mu_0)=\frac{\mu_0-\mu_1^*(\mu_0)-c}{-v_b-\mu_1^*(\mu_0)}$, $\lambda_1^*(\mu_0)=\frac{\mu_1^*(\mu_0)-c}{-v_b}$, and $\bar{\mu}_1^*(\mu_0)=\frac{-\mu_1^*(\mu_0)v_b}{\mu_1^*(\mu_0)-c}$ are continuous in μ_0.

We have shown in the proof of Lemma 7 that λ_0^* and μ_1^* strictly increase in μ_0 when μ_1^* is the interior solution. Now we consider the case when μ_1 is binding. When $\mu_1^*=\frac{c}{v_g}$, $\lambda_0^*=\frac{\mu_0-\mu_1^*-c}{-v_b-\mu_1^*}=\frac{\mu_0-\frac{c}{v_g}-c}{-v_b-\frac{c}{v_g}}$ strictly increases in μ_0. When $\mu_1^*=\frac{v_g\mu_0-c}{v_g-c}$, it is strictly increasing in μ_0 and $\lambda_0^*=\frac{\mu_0-\mu_1^*-c}{-v_b-\mu_1^*}=\frac{c}{v_g}$ is constant. Together with the continuity property we just established, we have shown that λ_0^* and μ_1^* (weakly) increase in μ_0. Thus, $\lambda_1^*=\frac{\mu_1^*-c}{-v_b}$ (weakly) increases in μ_0 and $\bar{\mu}_1^*=\frac{-\mu_1^*v_b}{\mu_1^*-c}$ (weakly) decreases in μ_0. □

According to the proof of Proposition 1 and Lemma 6, the sender does not provide information iff $\mu_0<\mu_{0,1}$ and provide information in one period if $\mu_{0,1}\leq\mu_0<\frac{2v_g-c}{(v_g)^2}c$. Thus, we just need to determine whether she provides information in one period or in both periods when $\mu_0\geq\frac{2v_g-c}{(v_g)^2}c$ by comparing the sender surplus of the optimal one-period strategy and the optimal sender surplus of the S_0 strategy.

(a) $c\leq v_g\lambda_0^{**}$: Define $\mu_{1,2}:=\inf\{\mu_0\geq\frac{2v_g-c}{(v_g)^2}c:\Pi_{S_0}(\mu_0)\geq\Pi_1(\mu_0)\}$. One can see that $\mu_{1,2}\in[\frac{2v_g-c}{(v_g)^2}c,\widehat{\mu_0})$ and $\Pi_{S_0}(\mu_{1,2})\geq\Pi_1(\mu_{1,2})$. According to Lemma 7,

$$
\begin{aligned}
\Pi_1(\mu_0)=&-K(\frac{\mu_0-c}{-v_b}\wedge\lambda_1^{**})+p(\frac{\mu_0-c}{-v_b}\wedge\lambda_1^{**})\\
\Pi_{S_0}(\mu_0)=&\max_{\mu_1}-K(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})+p\ \frac{\mu_0-\mu_1-c}{-v_b-\mu_1}+(1-\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})\left[-K(\frac{\mu_1-c}{-v_b})+\frac{(\mu_1-c)p}{-v_b}\right]\\
&s.t\ \mu_1\in\left[\frac{c}{v_g},\frac{v_g\mu_0-c}{v_g-c}\right]
\end{aligned}
$$

i) $\mu_{1,2} \geq c - v_b\lambda_1^{**}$: $\forall \mu_0 \in (\mu_{1,2}, \widehat{\mu_0}], \Pi_{S_0}(\mu_0) > \Pi_{S_0}(\mu_{1,2}) \geq \Pi_1(\mu_{1,2}) = \Pi_1(\mu_0)$.

ii) $\mu_{1,2} < c - v_b\lambda_1^{**}$: $\forall \mu_0 \in [\mu_{1,2}, c - v_b\lambda_1^{**})$, $\frac{d\Pi_1(\mu_0)}{d\mu_0} = K'(\frac{\mu_0 - c}{-v_b})\frac{1}{v_b} - \frac{p}{v_b}$.

A. $\mu_1(\mu_0) = \mu_1^u(\mu_0) := \frac{v_g\mu_0 - c}{v_g - c}$: In this case, $\lambda_0 = \frac{\mu_0 - \mu_1(\mu_0) - c}{-v_b - \mu_1(\mu_0)} = \frac{c}{v_g}$.
So, $\Pi_{S_0}(\mu_0) = -K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1 - \frac{c}{v_g})[-K(\frac{\mu_1^u(\mu_0) - c}{-v_b}) + \frac{(\mu_1^u(\mu_0) - c)p}{-v_b}]$.
For $\Delta > 0$ small enough, we have $\mu_0 + \delta < c - v_b\lambda_1^{**}$, $\forall \delta \in (0, \Delta)$. Consider $\mu_{0,\delta} = \mu_0 + \delta \in (\mu_0, \mu_0 + \Delta)$, we have $\Pi_{S_0}(\mu_{0,\delta}) \geq \underline{\Pi_{S_0}}(\mu_{0,\delta}) := -K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1 - \frac{c}{v_g})[-K(\frac{\mu_1^u(\mu_{0,\delta}) - c}{-v_b}) + \frac{(\mu_1^u(\mu_{0,\delta}) - c)p}{-v_b}]$. Noticing that $\Pi_{S_0}(\mu_0) = \underline{\Pi_{S_0}}(\mu_0)$, we have

$$\frac{d\Pi_{S_0}(\mu_0)}{d\mu_0} \geq \frac{d\underline{\Pi_{S_0}}(\mu_0)}{d\mu_0} = K'(\frac{\mu_1^u(\mu_0) - c}{-v_b})\frac{1}{v_b} - \frac{p}{v_b} > \frac{d\Pi_1(\mu_0)}{d\mu_0}$$

So, $\Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0), \forall \mu_0 \in [\mu_{1,2}, c - v_b\lambda_1^{**})$, and the inequality is strict when $\mu_0 > \mu_{1,2}$.

B. $\mu_1(\mu_0) < \mu_1^u(\mu_0)$: Let $\lambda_0 = \frac{\mu_0 - \mu_1(\mu_0) - c}{-v_b - \mu_1(\mu_0)}$. For $\Delta > 0$ small enough, we have $\mu_0 + \delta < c - v_b\lambda_1^{**}$ and $\mu_1(\mu_0) + \frac{\delta}{1 - \lambda_0} < \mu_1^u(\mu_0) < \mu_1^u(\mu_0 + \delta)$, $\forall \delta \in (0, \Delta)$. Consider $\mu_{0,\delta} = \mu_0 + \delta \in (\mu_0, \mu_0 + \Delta)$. Since $\frac{\mu_0 + \delta - (\mu_1^u(\mu_0) + \frac{\delta}{1 - \lambda_0}) - c}{-v_b - (\mu_1^u(\mu_0) + \frac{\delta}{1 - \lambda_0})} = \lambda_0$, we have $\Pi_{S_0}(\mu_{0,\delta}) \geq \widetilde{\Pi_{S_0}}(\mu_{0,\delta}) := -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)[-K(\frac{\mu_1(\mu_0) + \frac{\delta}{1 - \lambda_0} - c}{p}) + \frac{(\mu_1(\mu_0) + \frac{\delta}{1 - \lambda_0} - c)p}{-v_b}]$. Noticing that $\Pi_{S_0}(\mu_0) = \widetilde{\Pi_{S_0}}(\mu_0)$, we have

$$\begin{aligned} \frac{d\Pi_{S_0}(\mu_0)}{d\mu_0} \geq \frac{d\widetilde{\Pi_{S_0}}(\mu_0)}{d\mu_0} &= (1 - \lambda_0)\left[-\frac{1}{p}K'(\frac{\mu_1(\mu_0) - c}{p})\frac{1}{1 - \lambda_0} + \frac{1}{1 - \lambda_0}\right] \\ &= 1 - \frac{1}{p}K'(\frac{\mu_1(\mu_0) - c}{p}) \\ &\geq 1 - \frac{1}{p}K'(\frac{\mu_1^u(\mu_0) - c}{p}) \\ &> \frac{d\Pi_1(\mu_0)}{d\mu_0} \end{aligned}$$

So, $\Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0)$ and the inequality is strict when $\mu_0 > \mu_{1,2}$.

In sum, $\Pi_{S_0}(\mu_0) > \Pi_1(\mu_0)$, $\forall \mu_0 \in (\mu_{1,2}, \widehat{\mu_0}]$.

(b) $v_g\lambda_0^{**} < c < v_g\lambda_1^{**}$: The following lemma provides a closed-form solution to program (P_{2S_0}) when the search cost is intermediate:

Lemma 8. *Suppose $v_g\lambda_0^{**} < c < v_g\lambda_1^{**}$. $\lambda_0^* = \frac{c}{v_g}, \mu_1^* = \frac{v_g\mu_0 - c}{v_g - c}$ in the solution to (P_{2S_0}).*

Proof of Lemma 8. According to the proof of Proposition 10, λ_0^* is binding at the upper bound and increases in μ_0 in the solution to (P'_{2S_0}), for $c < v_g\lambda_1^{**}$. One can see that $\lambda_0^* = \frac{c}{v_g}$ and $\mu_1^* = c - v_b\lambda_1^{**}$ for μ_0 large enough in the solution to (P'_{2S_0}). Because $\lambda_0^* \geq \frac{c}{v_g}$, the only way for λ_0^* to be increasing in μ_0 is for it to always be $\frac{c}{v_g}$. Given $\lambda_0^* = \frac{c}{v_g}$, the optimal $\mu_1^* = \frac{v_g\mu_0 - c}{v_g - c}$ for (P'_{2S_0}). Lemma 6 shows that (P'_{2S_0}) is equivalent to (P_{2S_0}). So, $\lambda_0^* = \frac{c}{v_g}, \mu_1^* = \frac{v_g\mu_0 - c}{v_g - c}$ are also the solutions to (P_{2S_0}). $\square$

Define $\mu_{1,2} := \inf\{\mu_0 \geq \frac{2v_g - c}{v_g^2}c : \Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0)$ or S_+ strategy is feasible$\}$. Note that $\mu_{1,2} \leq \mu_{2,+}$. If the S_+ strategy is feasible $\forall \mu_0 > \mu_{1,2}$, the 1-period sender surplus will always be dominated by the 2-period sender surplus $\forall \mu_0 > \mu_{1,2}$, as the optimal S_+ strategy generates a strictly higher sender surplus than the optimal 1-period strategy. We now consider the case in which the S_+ strategy is not feasible for some $\mu_0 > \mu_{1,2}$, which implies that $\Pi_{S_0}(\mu_{1,2}) \geq \Pi_1(\mu_{1,2})$.

$$
\begin{aligned}
\Pi_1(\mu_0) =& -K(\frac{\mu_0 - c}{-v_b} \wedge \lambda_1^{**}) + p(\frac{\mu_0 - c}{-v_b} \wedge \lambda_1^{**}) \\
\Pi_{S_0}(\mu_0) =& -K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1 - \frac{c}{v_g})\left[-K(\frac{\mu_1^u(\mu_0) - c}{-v_b}) + \frac{(\mu_1^u(\mu_0) - c)p}{-v_b}\right]
\end{aligned}
$$

i) $\mu_{1,2} \geq c - v_b\lambda_1^{**}$: $\forall \mu_0 \in (\mu_{1,2}, \mu_{2,+}], \Pi_{S_0}(\mu_0) > \Pi_{S_0}(\mu_{1,2}) \geq \Pi_1(\mu_{1,2}) = \Pi_1(\mu_0)$.

ii) $\mu_{1,2} < c - v_b\lambda_1^{**}$: $\forall \mu_0 \in [\mu_{1,2}, c - v_b\lambda_1^{**}]$,

$$
\begin{aligned}
\frac{d\Pi_1(\mu_0)}{d\mu_0} =& K'(\frac{\mu_0 - c}{-v_b})\frac{1}{v_b} - \frac{p}{v_b} \\
\frac{d\Pi_{S_0}(\mu_0)}{d\mu_0} =& K'(\frac{\mu_1^u(\mu_0) - c}{-v_b})\frac{1}{v_b} - \frac{p}{v_b} > \frac{d\Pi_1(\mu_0)}{d\mu_0}
\end{aligned}
$$

So, $\Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0)$ and the inequality is strict when $\mu_0 > \mu_{1,2}$. $\Pi_{S_0}(c - v_b\lambda_1^{**}) > \Pi_1(c - v_b\lambda_1^{**})$. $\forall \mu_0 \in [c - v_b\lambda_1^{**}, \mu_{2,+}], \Pi_{S_0}(\mu_0) \geq \Pi_{S_0}(c - v_b\lambda_1^{**}) > \Pi_1(c - v_b\lambda_1^{**}) = \Pi_1(\mu_0)$.

In sum, $\forall \mu_0 \in (\mu_{1,2}, \mu_{2,+}],\ \Pi_{S_0}(\mu_0) > \Pi_1(\mu_0)$.

One can see that the optimal S_+ strategy always generates a strictly higher (and strictly positive) sender surplus than the optimal 1-period strategy. Therefore, the sender always provides information in both periods when the S_+ strategy is feasible.[16] By Lemma 5, the S_+ strategy is feasible iff $\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$ when $c < v_g\lambda_1^{**}$. Hence, together with the above results on the S_0 strategy, there exists $\mu_{1,2} \in [\frac{2v_g - c}{(v_g)^2}c, 2c - v_b\lambda_1^{**} \vee \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}]$ such that the sender does not provide information if $\mu_0 < \mu_{0,1}$, provides information in one period if $\mu_0 \in [\mu_{0,1}, \mu_{1,2})$, and provides information in both periods if $\mu_0 > \mu_{1,2}$. $\square$

Proof of Proposition 6.

(1) High Search Cost ($v_g\lambda_1^{**} \leq c < \widehat{c}$)

It has been shown in the proof of Proposition 5. Since the optimal strategy does not depend on the prior, the sender's payoff does not depend on the prior either.

[16] But the optimal 2-period strategy may be either S_+ or the S_0 strategy.

(2) Low Search Cost ($c \le \tilde{c} = v_g K'^{-1}\left[\frac{K(\lambda_1^{**})}{\lambda_1^{**}}\right]$)

$$\begin{aligned}
\text{F.O.C. of } (P_r) \Rightarrow & K'(\lambda_0^{**}) = K(\lambda_1^{**}) + p(1-\lambda_1^{**}) \\
\Rightarrow & \lambda_1^{**} K'(\lambda_0^{**}) - K(\lambda_1^{**}) = (1-\lambda_1^{**})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] > 0 \\
\Rightarrow & K'(\lambda_0^{**}) > \frac{K(\lambda_1^{**})}{\lambda_1^{**}} = K'(\frac{\tilde{c}}{v_g}) \\
\Rightarrow & \lambda_0^{**} > \frac{\tilde{c}}{v_g} \Rightarrow \tilde{c} < v_g\lambda_0^{**} < v_g\lambda_1^{**}
\end{aligned}$$

We now compare the optimal sender surplus between the solution to (P_{2S_+}) and the solution to (P_{2S_0}), and show that the optimal S_0 strategy is always preferred to the optimal S_+ strategy when both types of strategy are feasible.

Proposition 11. *Suppose $c \le \tilde{c}$ and $\mu_0 < \widehat{\mu_0}$. The sender uses the S_0 strategy when she provides information in both periods.*

Proof of Proposition 11. $\forall \mu_0 < \widehat{\mu_0}$ such that S_0 (S_+) strategy is feasible, denote the optimal sender surplus by $\Pi_{S_0}(\mu_0)$ ($\Pi_{S_+}(\mu_0)$).

i) $\mu_0' := \frac{(v_g-c)(c-v_b\lambda_1^{**})+c}{v_g} \le \mu_0 < \widehat{\mu_0}$: $\frac{(v_g-c)(c-v_b\lambda_1^{**})+c}{v_g} \le \mu_0 \Leftrightarrow \widehat{\mu_1}(\mu_0) \le c - v_b\lambda_1^{**}$. According to Proposition 9, $(\lambda_{0,S_+}, \mu_{1,S_+}) = (\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, c - v_b\lambda_1^{**})$ gives $\Pi_{S_+}(\mu_0)$. $\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c} \wedge c - v_b\lambda_1^{**}\right]$ in (P''_{2S_0}) and $\frac{(v_g-c)(c-v_b\lambda_1^{**})+c}{v_g} \le \mu_0 \Leftrightarrow \frac{v_g\mu_0-c}{v_g-c} \ge c - v_b\lambda_1^{**}$. Consider $(\lambda_0, \mu_1) = (\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}, c - v_b\lambda_1^{**})$, which satisfies the costraints in (P''_{2S_0}) and is identical to $(\lambda_{0,S_+}, \mu_{1,S_+})$. So, $\Pi_{S_0}(\mu_0) \ge \Pi_{S_+}(\mu_0)$.

ii) $\mu_0 < \mu_0'$: $\mu_0 < \mu_0' \Leftrightarrow \widehat{\mu_1}(\mu_0) > c - v_b\lambda_1^{**}$. According to Proposition 9, $(\lambda_{0,S_+}, \mu_{1,S_+}) = (\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, \widehat{\mu_1}(\mu_0))$ gives $\Pi_{S_+}(\mu_0)$. One can verify that $\frac{v_g\mu_0'-c}{v_g-c} = c - v_b\lambda_1^{**}$, $\frac{\mu_0'-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c} = \frac{c}{v_g} \le \lambda_0^{**}$. So, $\mu_0' \le \widehat{\mu_0}$, which implies that $\lambda_{0,S_+}(\mu_0') \le \lambda_0^{**}$.
Consider μ_0 such that both S_0 and S_+ strategies are feasible. Let $\mu_1^u(\mu_0) := \frac{v_g\mu_0-c}{v_g-c}$.

$$\begin{aligned}
\Pi_{S_+}(\mu_0) =& -K(\lambda_{0,S_+}(\mu_0)) + p\lambda_{0,S_+}(\mu_0) + (1-\lambda_{0,S_+}(\mu_0))[-K(\lambda_1^{**}) + p\lambda_1^{**}] \\
\Pi_{S_0}(\mu_0) \ge \underline{\Pi}_{S_0}(\mu_0) :=& -K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1-\frac{c}{v_g})\left[-K(\frac{\mu_1^u(\mu_0)-c}{-v_b}) + \frac{(\mu_1^u(\mu_0)-c)p}{-v_b}\right]
\end{aligned}$$

Lemma 9. $\frac{\mu_1^u(\mu_0)-c}{-v_b} \ge \lambda_{0,S_+}(\mu_0)$, $\forall \mu_0 < \mu_0'$.

Proof of Lemma 9. $\frac{d}{d\mu_0}[\frac{\mu_1^u(\mu_0)-c}{-v_b}] = \frac{v_g}{-v_b(v_g-c)}$, $\frac{d}{d\mu_0}[\lambda_{0,S_+}(\mu_0)] = \frac{1}{-v_b(1-\lambda_1^{**})-c}$.

$$\frac{d}{d\mu_0}[\frac{\mu_1^u(\mu_0)-c}{-v_b}] \le \frac{d}{d\mu_0}[\lambda_{0,S_+}(\mu_0)] \Leftrightarrow c(-2v_b-1) \le -v_bv_g\lambda_1^{**} \quad (*)$$

If $v_b \ge -1/2$, $(*)$ always holds. If $(-1 <)$ $v_b < -1/2$, we have that $\frac{-v_b}{-2v_b-1} \ge 1 \Rightarrow c \le v_g\lambda_1^{**} \le \frac{-v_bv_g}{-2v_b-1}\lambda_1^{**} \Rightarrow (*)$ also holds. So, $\frac{d}{d\mu_0}[\frac{\mu_1^u(\mu_0)-c}{-v_b}] \le \frac{d}{d\mu_0}[\lambda_{0,S_+}(\mu_0)]$, $\forall \mu_0 < \mu_0'$. Note that $\frac{\mu_1^u(\mu_0')-c}{-v_b} = \lambda_1^{**} > \lambda_0^{**} \ge \lambda_{0,S_+}(\mu_0')$. This concludes the proof. $\square$

Now we calculate the increasing rate of the sender surplus as a function of μ_0:

$$\frac{d\Pi_{S_+}(\mu_0)}{d\mu_0} = \frac{K'(\lambda_{0,S_+}(\mu_0)) - K(\lambda_1^{**}) + \frac{cp}{v_b}}{v_b(1-\lambda_1^{**})+c} - \frac{p}{v_b}$$

$$\frac{d\underline{\Pi}_{S_0}(\mu_0)}{d\mu_0} := \frac{K'(\frac{\mu_1^u(\mu_0)-c}{-v_b})}{v_b} - \frac{p}{v_b}$$

$$\frac{d\Pi_{S_+}(\mu_0)}{d\mu_0} \geq \frac{d\underline{\Pi}_{S_0}(\mu_0)}{d\mu_0} \Leftrightarrow \frac{K'(\frac{\mu_1^u(\mu_0)-c}{-v_b})}{-v_b} + \frac{K(\lambda_1^{**}) - \frac{cp}{v_b}}{-v_b(1-\lambda_1^{**})-c} \geq \frac{K'(\lambda_{0,S_+}(\mu_0))}{-v_b(1-\lambda_1^{**})-c} \qquad (\star)$$

$$c \leq \tilde{c} = v_g K'^{-1}\left[\frac{K(\lambda_1^{**})}{\lambda_1^{**}}\right] \Leftrightarrow K(\lambda_1^{**}) \geq \lambda_1^{**} K'(\frac{c}{v_g})$$

$$\Rightarrow -v_b\left[K(\lambda_1^{**}) - \frac{cp}{v_b}\right] \geq (c - v_b\lambda_1^{**})K'(\frac{c}{v_g}) \left(c \leq v_g\lambda_1^{**} \Rightarrow K'(\frac{c}{v_g}) \leq K'(\lambda_1^{**}) = p\right)$$

$$\Rightarrow -v_b\left[K(\lambda_1^{**}) - \frac{cp}{v_b}\right] \geq (c - v_b\lambda_1^{**})K'(\lambda_{0,S+}(\mu_0)) \left(\lambda_{0,S+}(\mu_0) < \lambda_{0,S+}(\mu_0') = \frac{c}{v_g}\right)$$

$$\Leftrightarrow \frac{K'(\lambda_{0,S+}(\mu_0))}{-v_b} + \frac{K(\lambda_1^{**}) - \frac{cp}{v_b}}{-v_b(1-\lambda_1^{**})-c} \geq \frac{K'(\lambda_{0,S_+}(\mu_0))}{-v_b(1-\lambda_1^{**})-c}$$

$$\overset{Lemma\,9}{\Rightarrow} \frac{K'(\frac{\mu_1^u(\mu_0)-c}{-v_b})}{-v_b} + \frac{K(\lambda_1^{**}) - \frac{cp}{v_b}}{-v_b(1-\lambda_1^{**})-c} \geq \frac{K'(\lambda_{0,S_+}(\mu_0))}{-v_b(1-\lambda_1^{**})-c}$$

$$\overset{(\star)}{\Leftrightarrow} \frac{d\Pi_{S_+}(\mu_0)}{d\mu_0} \geq \frac{d\underline{\Pi}_{S_0}(\mu_0)}{d\mu_0}$$

One can verify that $\lambda_{0,S_+}(\mu_0') = \frac{c}{v_g}$, $\frac{\mu_1^u(\mu_0')-c}{-v_b} = \lambda_1^{**}$. So, $\Pi_{S_+}(\mu_0') = \underline{\Pi}_{S_0}(\mu_0')$. Therefore, $\underline{\Pi}_{S_0} \geq \Pi_{S_+}(\mu_0)$. $\Pi_{S_0}(\mu_0) \geq \underline{\Pi}_{S_0} \Rightarrow \Pi_{S_0}(\mu_0) \geq \Pi_{S_+}(\mu_0)$.

Now we show that the S_0 strategy is feasible whenever the S_+ strategy is feasible, which concludes the proof of Proposition 11.

Lemma 10. *Suppose $c < v_g\lambda_1^{**}$. For any μ_0 such that the S_+ strategy is feasible, the S_0 strategy is also feasible.*

Proof of Lemma 10. Suppose there exists μ_0 such that the S_+ strategy is feasible while the S_0 strategy is not feasible. Then, $2c - v_b\lambda_1^{**} < \frac{2v_g-c}{(v_g)^2}$ and $\frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c} < \frac{2v_g-c}{(v_g)^2}$, which is equivalent to $\lambda_1^{**} < \frac{c^2}{v_b(v_g)^2} + \frac{2c}{v_g}$ and $c > v_g\lambda_1^{**}$, which is not possible as we assumed that $c < v_g\lambda_1^{**}$. □

□

Proposition 11 tells us that we can limit our attention to the S_0 strategy when $c \leq \tilde{c}$. Proposition 10 has characterized the optimal S_0 strategy, which implies that the sender's payoff strictly increases in the prior belief.

□